Practical Industry in CG Ikhlas SDN BHD

By
LAI CHUN KIT

Dissertation submitted in partial fulfilment of the
requirement for the award of the degree of
Bachelor of Science in Financial Mathematics (with Honours)

FACULTY OF COMPUTER SCIENCE AND MATHEMATICS
UNIVERSITI MALAYSIA TERENGGANU
2026

DISSERTATION CONFIRMATION AND APPROVAL

This is acknowledged and confirmed that the dissertation entitled: *Practical Industry in CG Ikhlas SDN BHD* by Lai Chun Kit, Matric No: S65947 have been checked and all the suggested corrections have been done. The dissertation is submitted to the Faculty of Computer Science and Mathematics, Universiti Malaysia Terengganu in partial fulfilment of the requirements for the award of the degree of Bachelor of Science in Financial Mathematics with Honours.

Authorized by:

…….. Date: ……...
Main Supervisor
Name:
Official Stamp:

…….. Date: ……...
Co-Supervisor (If any)
Name:
Official Stamp:

…….. Date: ……...
Practical Industry's Coordinator
Bachelor of Science in Financial Mathematic with Honours
Name: Puan Siti Madhihah binti Abdul Malik
Official Stamp:

DECLARATION

I hereby declare that this dissertation is the result of my own research except as cited in the references.

Signature : ……
Name : Lai Chun Kit
Matric No. : S65947
Date : 20/1/2026

ACKNOWLEDGEMENTS

First and foremost, I would like to express my utmost appreciation towards CG Ikhlas Sdn Bhd (CGI) for offering me the chance to complete my industrial training in this prestigious company. This has been instrumental in closing the loop between what I know academically and how to apply it in a practical sense within data analytics.

I could like to express my heartfelt appreciation to Ms. Elise Chong, who is my industrial supervisor for her guidance, trust and advice during these internship period. I am truly grateful for the freedom she gave me to head knives in high-impact initiatives such as; The Pricing Intelligence Engine (PIE), The Shopee Product Strategy Analyzer, and dozens of custom-built automation data collection tools. She has given me the trust to advocate a move from hand on keyboard data entry to sophisticated & python automation, where my management skills would not be as sharp.

And thank you also to everyone at CGI, especially on the Marketing and Technical sides. I appreciate their openness to take tools I built and diligently incorporate them in their day to day. By using these tools, optimizing for the best efficiency and productivity we were going to take fantastic results from 11.11 Campaign monitoring. Especially, thanks to data-based strategies permitted by these tools, sales significantly have improved from the level of RM70,000 in 2024 to that of RM120,000 in 2025.

An appreciation also to my supervisor from Universiti Malaysia Terengganu (UMT), Dr. Auni Aslah Bin Mat Daud, thanks for your never degree guidance and monitoring towards the writing of this thesis. Special thanks to his comments provided in the site visit, which didn't go unnoticed and have been helpful in terms of framing

my project presentations right, allowing me to steered my knowledge in the right directions.

To my family and friends, I couldn’t have done this without your love and support. It's been a life changing job training where I have learned new web scraping, database and system architecture skills which will be invaluable to my future career.

PRACTICAL INDUSTRY IN CG IKHLAS SDN BHD

EXECUTIVE SUMMARY

During my 24-week internship at CG Ikhlas Sdn Bhd, which is widely recognized for its retail brand i-Store and the official platform ihaveit.com.my, I served as a data analyst intern. The company is a prominent Malaysian IT solutions provider and an authorized e-tailer based in Subang Jaya. CGI focuses on the distribution of high-performance computing hardware. It is a certified partner for Asus and the Republic of Gamers. My main objective from August 2025 to January 2026, is modernising the company's retail operations by trying to replace manual and time-consuming processes with data automation and strategic performance tracking tools.

I specifically developing the Shopee and Lazada Product Strategy Analyzers for the 11.11 Mega Campaign. These tools were designed to split up large number of products into actionable strategic groups. The Engagement Remediation section is mainly identified products that attracted high traffic but suffered from low conversion rates. It enabled the marketing team to rapidly identify which listings had issues and rerouted ad spend much more efficiently. Simultaneously, I utilized the Scale Winner section to highlight best-selling products that I had then prioritized for greater ad spend to drive sales volume. As a result of its data driven approach, performance grew rapidly with campaign sales increasing 70% from RM70,000 in 2024 to in excess of RM120,000 in 2025.

After the 11.11 campaign, I moved onto my main project, which was the Price Intelligence Hub. This project began with an experimentation phase in which I built a semi-automated Chrome extension to study how to scrape prices information directly

from e-commerce pages. It eventually matured into a full stack automation system, including backend services and a single page web application that integrated the manual review and approval workflow. This hub now serves as the nerve center for operations and gives the company market details so swiftly and accurately that it was previously impossible to achieve manually.

To ensure that the hub sounded smart and reliable when it operated, I built two custom automated collection tools. The first is a seller name collector that is designed to glide right past any security measures and captchas without ever needing a user to log in. With rotating IP VPNs, I protected the system from having seller data missed and I protected the system from being banned by the platform. The second is a price scraper that use to get detailed information on competitors. To protect against unauthorized access to the company's most important accounts, I connected this tool to burner accounts, not official accounts, and then used static VPN IPs. By strategically combining both methods, it enables to maximize data collection while completely eliminating the possibility of our main accounts being frozen/indefinitely banned from the platform.

In summary, this industrial training was an essential link between technical data engineering and organisational impact. I played a key role in the historic growth of the company by developing tools to optimize ad spend through remediation and scaling strategies while growing my professional skill set in Python, SQL, and business intelligence. Through this practical exposure, I have gained the technical expertise and industry knowledge that has prepared me to be a successful data analyst and innovate the company from a manual reporting organisation to an agile, automated, and a high growth organisation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviations	
AJAX	Asynchronous JavaScript and XML
API	Application Programming Interface
B2B	Business-to-Business
B2C	Business-to-Consumer
BAU	Business as Usual
CGI	CG Ikhlas Sdn Bhd
CI	Competitive Intelligence
COM	Component Object Model
CPC	Cost Per Click
CSS	Cascading Style Sheets
CTR	Click-Through Rate
DAX	Data Analysis Expressions
DOM	Document Object Model
DSS	Decision Support System
ERP	Enterprise Resource Planning

FK	Foreign Key
GPU	Graphics Processing Unit
HTML	HyperText Markup Language
IO	Internal Order
IT	Information Technology
JS	JavaScript
JSON	JavaScript Object Notation
PIE	Pricing Intelligence Engine
PO	Purchase Order
POE	Proof of Execution
RAM	Random Access Memory
RMA	Return Merchandise Authorization
ROAS	Return on Ad Spend
ROG	Republic of Gamers
RRP	Recommended Retail Price
SDLC	Software Development Life Cycle
SKU	Stock-Keeping Unit
SME	Small and Medium-sized Enterprise
SOA	Service-Oriented Architecture
SQL	Structured Query Language

SSD	Solid State Drive
UI	User Interface
UMT	Universiti Malaysia Terengganu
WAF	Web Application Firewall
YOY	Year-over-Year

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION TO THE ORGANIZATION

1.1 Introduction

This chapter gives a review of the company where industrial training was taken at namely CG Ikhlas Sdn Bhd (CGI). It covers the goals of the industrial training, a description of the company profile and history background (company organizational structure such as vision, mission & objective), organisation chart and its departments core functions.

1.2 Objectives of Industrial Training

Industrial training is an obligatory part of the academic syllabus for the purpose to provide students with their initial realization on how they can work in industry.

The aims of this programme are:

i. To provide students with industrial exposure by learning to work in a fast-paced e-commerce and Information Technology (IT) environment.
ii. To use theoretical insights, especially in data analysis, programming and statistics to solve business problems.
iii. To establish and maintain professional connections with industry professionals for future career opportunities.
iv. To develop technical hard skills in Python programming, Excel, SQL database management and applied mathematics, alongside fundamental soft skills in communication and complex problem solving.

1.3 Organization Information

1.3.1 Organization Name and Address

Name: CG Ikhlas Sdn Bhd

Address: No 34 & 36, Jalan Industri USJ 1/11, Taman Perindustrian Usj 1, 47600 Subang Jaya, Selangor

1.3.2 Organization Background

CG Ikhlas Sdn Bhd (CGI) is a reputable company in Malaysian IT retail market, which is widely recognized by the public through its store-based retailer, i-Store and an extensive online portal, ihaveit. com. my. Tucked centrally in the heart of Subang Jaya, the company is a go-to store for consumers, gamers and enthusiasts who are looking for highend computing hardware.

To the client, CGI is the "One-Stop IT Hub" giving that peace of mind. The shop is a Multi-Brand Authorized Dealer, so it retails authentic products from worldwide tech leaders. Though they're best known as a Gold Partner for ASUS and Republic of Gamers (ROG), their product list also includes full ranges from Lenovo, MSI, HP, Dell, Cooler Master and etc. Whether you're that customer who stops by the showroom to play around with a gaming laptop or someone who's ordering components via its website, its official manufacturer warranties ensure you get the best shopping experience.

1.3.3 History of the Establishment of the CG Ikhlas SDN BHD

CGI was created in response to an important gap in the market, namely that for a reliable reseller of IT equipment from whom customers could purchase genuine hardware with confidence as opposed to grey imports or counterfeit products. From a small IT store in Subang Jaya, it has now evolved into a full-fledged omnichannel retailer.

The company expanded beyond that physical shop over the years to follow a need for shopping online. It has a strong online presence today with well-rated official stores on Shopee and Lazada, alongside its own website. That progression has enabled the company to serve customers across the country, online with the same knowledge and diverse array of products that they would find in any one of its showrooms.

1.3.4 Vision, Mission, Motto, Slogan, and Organizational Objectives

i. Vision

To become the preferred technology destination for Malaysian consumers, synonymous with quality, variety and customer service excellence.

ii. Mission

a) Deliver confidence to customers that all product sold is 100% authentic and will come with the local warranty.
b) Provide a higher level of shopping experience by proactively establishing technical service, fast logistics and thoughtful after-sales service.
c) To deliver the newest technology at a price that's competitive with other outlets, while offering a variety of ways to shop.

iii. Objective

a) To provide the widest variety of IT hardware on the market for an average customer to easily be able to find everything they need from gaming laptops in one place, all the way down to office equipment.
b) Keep up the good work when it comes to positive customer satisfaction score on platforms like Shopee, Lazada, Tiktok.
c) Consolidate relationship with brand principles (such as ASUS & Lenovo), due to which customer gets the 1st priority on new product launches and exclusive offers.

1.3.5 Organization Logo

Figure 1-1 The i-Store logo, which is widely recognized by customers in Subang Jaya and online shoppers as a symbol of trust and authentic IT hardware.

1.4 Organizational Activities

A SME with main business in ecommerce, CG Ikhlas Sdn Bhd is running a light and flexible management style. While the business does maintain a showroom, it acts primarily as fulfillment center and not in the traditional retail sense. This means that their day to day is focused around everything from e-commerce and logistics, to backend support. With the fast-moving environment of the business, functional responsibilities do indeed overlap and hence close interdisciplinary relationship between the following core departments is essential:

1.4.1 General Management

The management of the organization is based on General Management that has the strategic and directional control. This department is responsible for the maintenance of relationship senior with brand principals (e.g. ASUS, LENOVO) as well as setting long term business targets that provides guiding light to all other functions.

1.4.2 Marketing and Creative Media

The strategic center of the company's B2C business, as online retailer takes place in the Marketing and Creative Media department. The latter department plans the campaign calendar around high-traffic e-commerce events, like Payday Sales and

monthly “double-date” days. Importantly, this unit also partners with the Graphic Design sub-unit. All promotional banners, product assets, and store decor to push higher CTR (click through rate) & conversions within digital store fronts are created from this team.

1.4.3 Finance, Administration, and Human Resources

Headcount department provides organizational framework, structural solidity, and financial transparency. One is the administrative and accounting functions, which includes the strict verification of bills and auditing of transactions, particularly the margins within the competitive IT sector. At the same time, the Human Resources function is managing workforce planning and recruitment to ensure the operational capacity of the company can grow in line with the spikes in labour demand generated by peak sales periods.

1.4.4 Data Analytics and Business Intelligence

Data Analytics and Business Intelligence unit served as the modernization arm of the organization. This unit primarily focuses on converting manual workflows into automated, data-driven processes. This unit serves as the "Decision Support System"(DSS) for the Marketing team to spend more on Ads and Management to make pricing decisions on-the-fly by interpreting large-scale market data and pricing trends of competitive brands.

1.4.5 Sale Operations (Inside Sales)

Sale Operations unit is the first line of contact for digital retail outlets on platforms like Shopee and Lazada. On this team, they mainly monitor and verify incoming online orders and stock in real-time. The idea is to validate whether or not a digital transaction can fulfil against physical inventory levels, before the fulfillment of the transaction, in order to generating accurate seller ratings and reducing order cancellations.

1.4.6 Business Development (Outside Sales)

Unique from the internal advertising evaluate, the Business Development and Outside Sales unit pursues B2B (Business-to-Business) expansion and corporate sales. This department handles all the tedious transactional governance of large-scale transactions and processing of IO (Internal Order) and PO (Purchase Order). The distinction lets the firm target extended company contracts without interfering with the daily e-commerce business cycle.

1.4.7 Inventory, Technical Support, and Logistics

This department oversees the entire physical product lifecycle and "last-mile" execution. The Inventory and Technical Support sub-unit is responsible for replenishing stocks and offers value-added services like hardware upgrades (i.e., RAM/SSD installations) and system troubleshooting. The Logistics and Fulfillment team did all the heavy lifting of secure focused pick, pack and ship. This unit also takes care of warehouse arrivals for incoming supplier shipments, ensuring that bigger stocks are arranged in a required table whilst synced to the digital inventory.

1.5 Organizational Chart

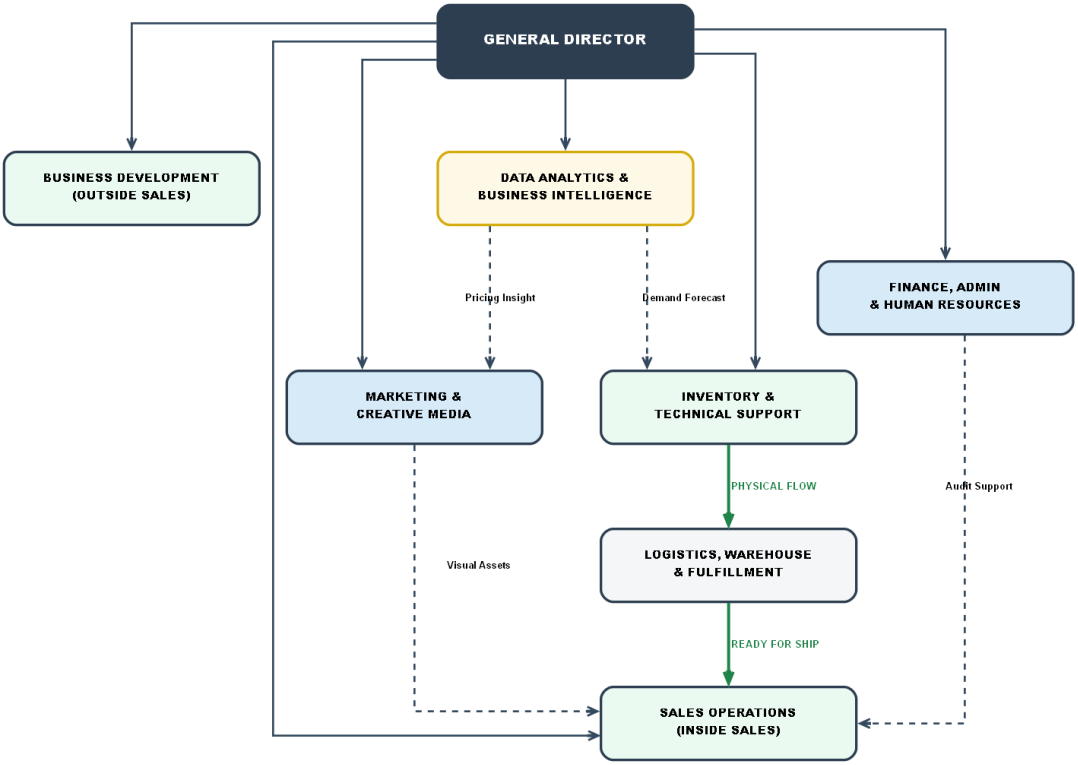


Figure 1-2 Functional Organizational Structure of CG Ikhlas Sdn Bhd

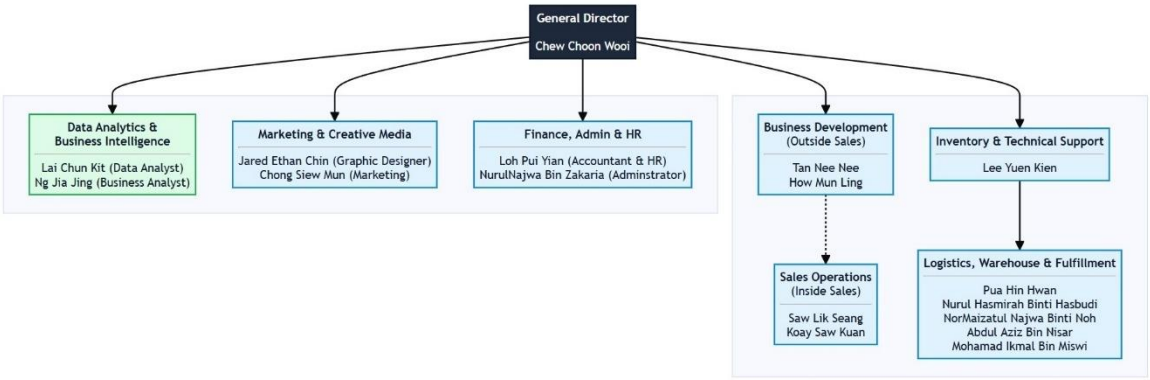


Figure 1-3 Organizational Chart of CG Ikhlas Sdn Bhd

1.6 Organizational Functions

Being one of the forefronts of the Malaysian digital economy, CG Ikhlas Sdn Bhd has multiple essential services that connects international tech firms with end-users:

i. Authorized distribution channel

The organisation acts as a gateway to Principals around the world including ASUS, Lenovo and MSI. This is not to prevent imports as such, it is more to ensure a genuine and non-unauthorized hardware is available in the Malaysian markets. All the products are sourced from OEM (original equipment manufacturer) channels to assure quality which is the leading cause of concern while purchasing products online.

ii. E-Commerce Fulfillment Hub

The organisation is a digital retail fulfilment centre of great volume. It handles the entire lifecycle of orders made online from Shopee, Lazada and its own website. Such services entails maintaining stock, secure packing and logistics for immediate delivery to customers nationwide.

iii. Technical Value Add

Unlike"box-move" resellers, CGI is a technical service provider. The company distinguishes itself by providing pre-delivery customization (upgrades to RAM and SSD, for example) while acting as an intermediary for warranty claims (RMA). And this feature allows users to get the hardware best suited for them right out of the box.

iv. Market Intelligence and Pricing Strategy

Following the implementation of business intelligent systems, the company itself is a data-driven organisation. It constantly tracks market trends and competitor prices to maintain dynamic pricing strategies. It is this function

through which the company lead the market by being a price leader and ensuring profit at an operational level.

1.7 Location Map and Building Photos

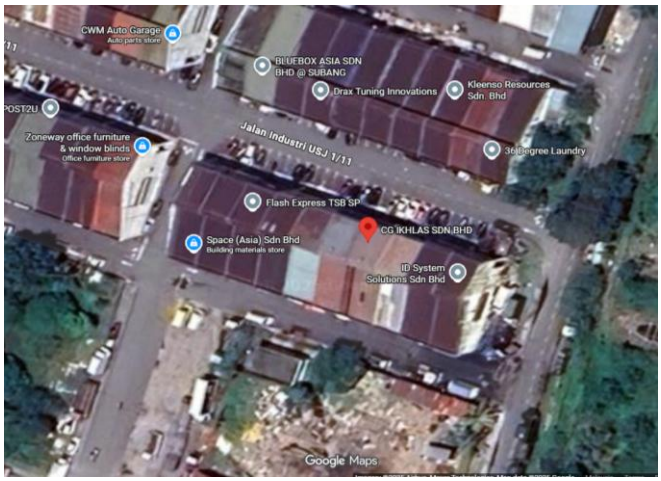


Figure 1-4 Chart Location Office CG Ikhlas SDN BHD

Figure 1-5 The showroom of CGI

Figure 1-6 The front building view of CG Ikhlas SDN BHD

Figure 1-7 The office front door entry view of marketing department

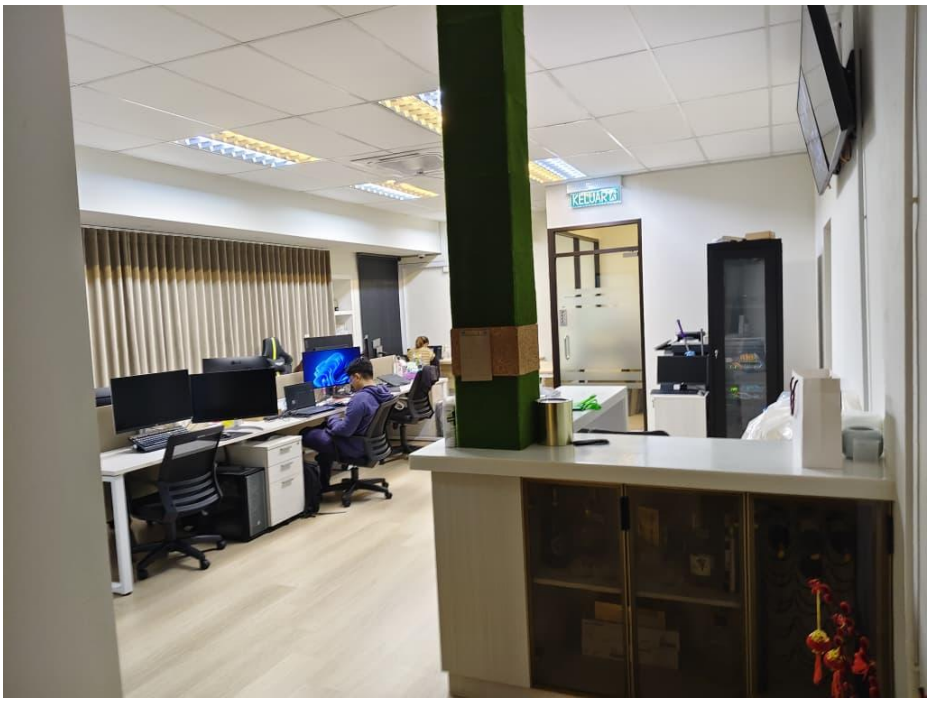


Figure 1-8 The office view of marketing department

CHAPTER 2

MINI PROJECT

2.1 Introduction

Chapter 2 provides a detailed technical exposition of the Price Intelligence Hub V3, a system developed to serve as the primary source of truth for market intelligence and competitive pricing. The project was conceptualized to modernize traditional retail operations by transitioning from labor-intensive manual competitor monitoring to a fully automated data acquisition framework. However, as the company grew and its product catalogue increased, the volume and complexity of the work became too much for the previous manual workflows, creating a need for an efficient digital solution that could scale and offer market intelligence within a report's turnaround time.

In order to maximize performance and maintainabilty, the system is built using a Service-Oriented Architecture (SOA) via a loosely coupled Controller-Worker design pattern. It combines four basic functional pillars; the Web Controller that handles session management and serving the UI, the Collector Engine that handles automated data scraping, the strategy engine that handles automated Excel sync over the COM interface, and the Intelligence Engine that handles data normalization using Machine Learning. With the introduction of Price Intelligence Hub V3, these components are centralized to provide the organization with a quick market view and a differentiated competitive edge that was not possible with manual processes. The upcoming sections describe the technical approach, the specific "Stealth Scraping" algorithms used and the overall results seen as a consequence of the system.

2.2 Project Background and Objectives

2.2.1 Project Background

Leading up to the build of Price Intelligence Hub V3, CG Ikhlas Sdn Bhd struggled with operational hurdles that impact on competing for pricing across their e-commerce platforms. With a wide range of products spanning in thousands of Stock-Keeping Units (SKUs) on the marketplaces like Shopee, Lazada this current workflow was manual competitor tracking intensive. This manual process was not only time consuming and tedious, but errors were made by human hands and “blind spots” occurred where the company did not actually realize it was pricing its product higher than competitors in the market.

And there was no real-time data, so pricing actions were more reactive. The marketing team would often price large campaigns on gut feel or outdated reports, so the company was leaving money and profit margin by chance. To address these inefficiencies the firm required a unified, automated system capable of capturing, analyzing and reporting on market data quickly and accurately. That requirement inspired the creation of a custom “Price Intelligence Hub” that serves as the single source of truth for pricing decisions.

2.2.2 Objectives

The project’s destination was to revolutionise the company’s retail by innovation. The specific objectives were:

i. To create a "headless" automated data collector engine, which would continuously scrape high volume pricing data from competitors' store thereby doing away with the cumbersome manual entering of Pricing Data.
ii. To develop a complete human imitation layer to implement the "Ghost Mouse" protocol, naturalized keystroke dynamics and heuristic scrolling patterns, in order to effectively bypass behavioral biometric security mechanisms.

iii. To develop a live scoring engine in Excel through COM interfaces, bringing the technical database closer to the way marketing staff work and making it possible for non-technical users to make data-driven decisions.

iv. To create a real-time, graphic representation of market positioning and "Trust Tiers" in a single intelligence dashboard for instant recognition and correction of competitive pricing gaps.

2.2.3 Literature Review

The evolution of the Price Intelligence Hub V3 has been built on academic research, and industry-dominant concepts of data mining, IT security and strategic market placement. In formal terms, Web scraping is an automated software technique to extract large amounts of information from websites and then convert it into a structured data for business analysis purposes. According to Glez-Peña et al. (2014), the effectiveness of information retrieval in contemporary e-commerce heavily relies on a system's capability to handle and parse complex DOMs. Since modern platforms like Shopee heavily rely on JavaScript dynamics, testing and automation frameworks such as Selenium are required to obtain live pricing information that conventional static scrapers cannot retrieve.

Such systems' architectural robustness is typically tackled from the perspective of software engineering patterns. Ferrara et al. (2014) claim that web information extraction has become one of those changing forms, and its technologies involve the development of solid methodological tools to address the heterogeneity of web sources and usually follows decoupled architectures to guarantee scalability. The Price Intelligence Hub V3 removes the data acquisition "worker" from the central web controller in a Service-Oriented Architecture (SOA) allowing for significant numbers of concurrent SKU requests to be executed without suffering any system degradation. This is so that we can keep the inflow of data steady even during times of heavy platform traffic, and continue to have a live view into competitive landscape for the company.

The argument of strategic positioning on the marketplace is empowered by Competitive Intelligence (CI) theory that defined it as a set of methods and techniques for an ethical and systematic information gathering and processing about competitor activities, in the service of reaching competitive advantages. According to Bose (2008) CI acts as a strategic weapon of intelligence analysis that would enable companies to spot market changes and competitors' threats earlier than they do it on financial record. Automating the acquisition of public market data, the company is able to turn external environmental signals into proprietary internal intelligence which enables its members to be more timely and informed when making business decisions in digital marketplace nowadays.

Dynamic pricing — where the prices of products are adjusted in real time based on market demand and competitive positioning — is considered a matter of survival in cutthroat e-commerce. Kannan and Kopalle (2001) observe that the capability of companies to watch for and respond almost immediately in price changes through the internet implies that consumers should not automatically expect to receive a consumer surplus. Specifically, the Strategy Engine we built in helps to enact some of these theories by providing pricing updates at the SKU level and enabling "Trust Tiers" where you don't even have to look at all of your SKUs but can be at a competitive state.

While e-commerce websites continue to improve bot-detection technology, behavioral biometrics are emerging as a key enabler for stealth data acquisition. Feher et al. (2012) show that bots can be detected and differentiated from human users organically by analyzing mouse dynamics, especially movement trace, velocity and clicks. To alleviate this, the project leverages a human emulation layer through "Ghost Mouse" algorithm. This algorithm uses non-linear Cubic Bezier curves to simulate the jitter and acceleration of a human hand. Additionally, the system replicates naturalized keystroke behavior of random latent input characters and heuristic scrolling muscle memory logic to emulate human reading habits which can effectively evade behavior based security measures such as CAPTCHA.

An important technical bottleneck in market monitoring is entity resolution; ie, correctly determining which disparate text strings by different sellers correspond to the same physical good. Getoor and Machanavajjhala (2012) claim that heuristics-based matching systems are required for ensuring database integrity and fighting duplicates in massive databases with the use of brand filtering and specification validation. The system's Four-Tier Matching Logic filters out unrelated competitor listings, or what it calls "false positive" matches which often occur with simpler keyword-based search algorithms.

Finally, for SMEs, the process of moving from intuition to data-driven decisions is a necessity in the long run if it has any hope of becoming competitive. Turban et al. (2014) consider Decision Support Systems (DSS) as automated systems supporting managers in solving semi-structured problems by making available analytical models and data. This trend is facilitated by the Resource-Based View (RBV) of the firm, in which Barney (1991) suggests that a technology held to be valuable, rare and non-imitable can lead to sustained competitive advantage. With the use of a Python intelligence backend combined with an Excel COM interface, The Price Intelligence Hub is effectively a DSS which will provide an organization with their own unique power-user resource.

2.3 Methodology

The implementation and live operation of the Price Intelligence Hub V3 which is shown in the Figure 2-1 is based on an extensive five phase technical pipeline connecting manual e-commerce administration work-flows with high-volume automated business intelligence. This process follows the full lifecycle of data, from ingestion and staging with multiple sources through to a complicated Excel-based strategy engine to fully automated market acquisition and trust-weighted analytic reporting.

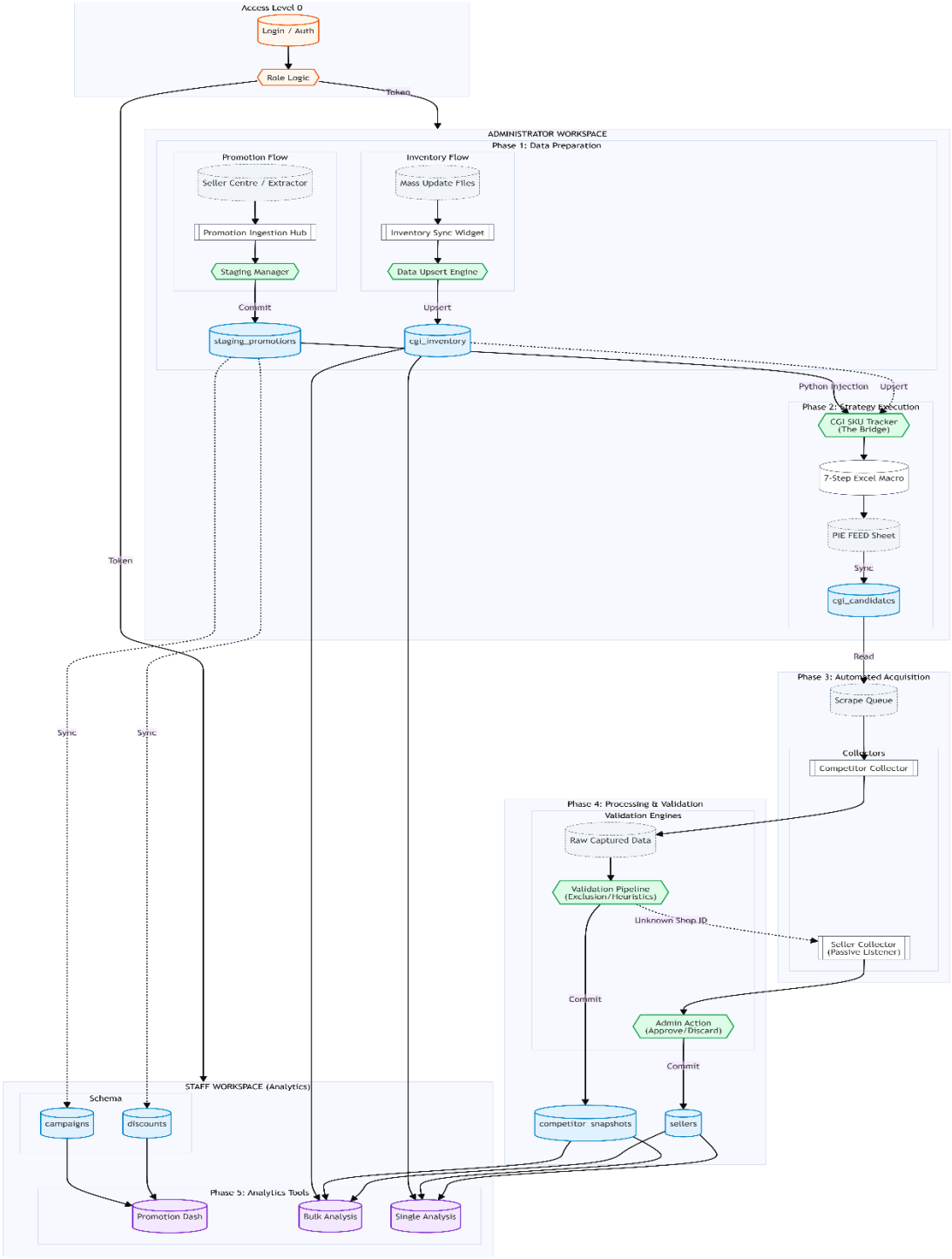


Figure 2-1 Price Intelligence Hub V3 System Architecture. This shows the complete 5-phase Data Pipeline flow transition in which all critical steps are completed for every stage of activity ends successfully, from Data Preparation (Phase 1) and Strategy Execution (Phase 2) done in Administrator Workspace through to Automated Acquisition agents (Phase 3) and Validation Pipelines (Phase 4)--via the final Analytics & Reporting Dashboard (Phase 5) used by staff even when all resources have been redistributed.

2.3.1 Phase 1: Data Preparation and Multi-Source Ingestion

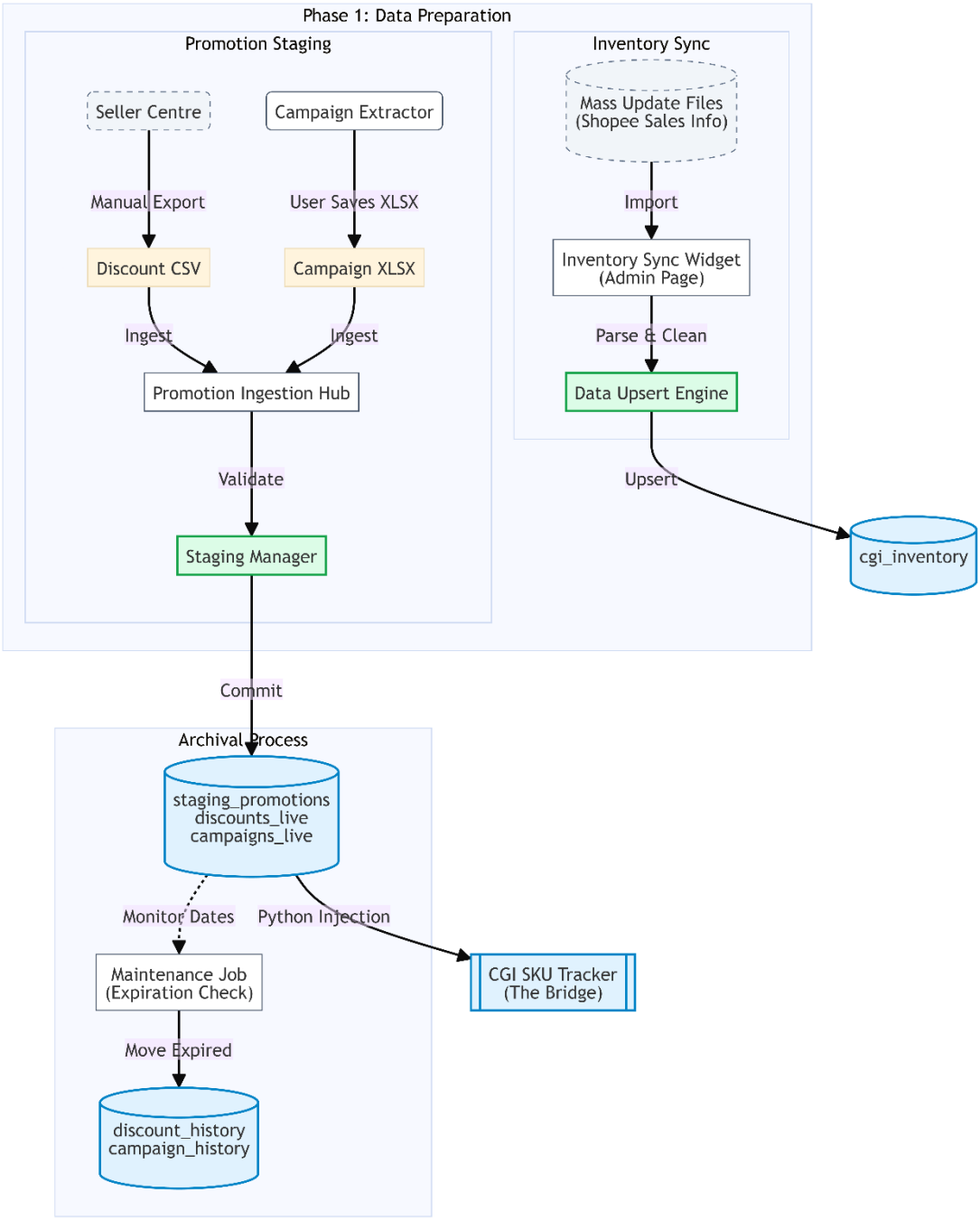


Figure 2-2 Phase 1 – Data Preparation and Multi-Source Ingestion. The workflow for ingesting raw inventory and promotion data from heterogeneous sources (Excel, CSV). The diagram shows the staging logic including the "Inventory Sync Widget" for upserting records and the "Archival Process" (Maintenance Job) that automatically migrates expired promotions to historical storage to maintain database performance.

The foundational stage of the methodology, shown in Figure 2-2, is about linking those records that dwell inside the system with information out in the volatile marketplace. Instead of forcing analysts to spend precious time and energy on manual data fetching and data cleansing, this phase electrically generates two separate input data streams: basic inventory folder and real-time multi-platform promotional inputs.

A. Inventory Synchronization and Database Upsert

Step one is to manually download the Mass Update Sales Info files (Parts 1 and 2) from the Shopee Seller Centre. The data is then imported into Structured Reports via an Inventory Sync Widget on the administrator dashboard, which uses a powerful Data Upsert Engine to interface with a live system. This engine carries out a UPSERT operation (Update-if-exists, Insert-if-new) on the *cgi_inventory* table in the database, so that critical fields like stock levels, listing prices, and SKU status are refreshed. At the same time, the system carries out a Python-driven injection to push this new data into the *Stock_List* sheets of the master Excel workbooks, thereby bypassing manual data entry and ensuring consistency across all platforms.

B. Multi-Stream Promotion Ingestion

To accommodate the complexity of e-commerce digital retail marketing, for this first release (Phase 1) a split-source ingestion logic is used to ingest promotional data:

i. Discount Promotions: Regular seller promotion discounts are manually exported as CSV from the Seller Centre and saved in a dedicated local directory.
ii. Campaign Promotions: High frequency episodic events, e.g., 11.11, 12.12 are captured through a custom Campaign Extractor Extension. This extension scrapes unstructured campaign data and returns it into the browser DOM, validating against excel sheet.

C. The Promotion Staging Hub and Validation Gate

Both promotional streams are funneled into the Promotion Ingestion Hub, where undergo a mandatory validation process within a "Staging Area" (*staging_promotions* table). This staging layer serves as a quality control gate where the administrator reviews file integrity, price formats, and promotion timelines, specifically the Start and End dates, before the data is committed to the live environment. Once verified, a "Commit" trigger moves the data into the *discounts_live* and *campaigns_live* tables, while an automated maintenance job monitors these records to archive expired entries into historical tables for future auditing.

D. Operational Loop Closure: The Data Bridge

The final technical operation of Phase 1 is the CGI SKU Tracker (Discount & Campaign).xlsx. The system will using a specialized Python injection script to extract the validated records from the live promotion tables and populates this "Bridge" file. This SKU tracker later on will be acting as the centralized artifact of Phase 1, and push to the structured data source that the Price Comparison.xlsm workbook via Power Query to calculate final strategic prices in the subsequent methodology phases.

2.3.2 Phase 2: Strategy Execution (The Formula Engine)

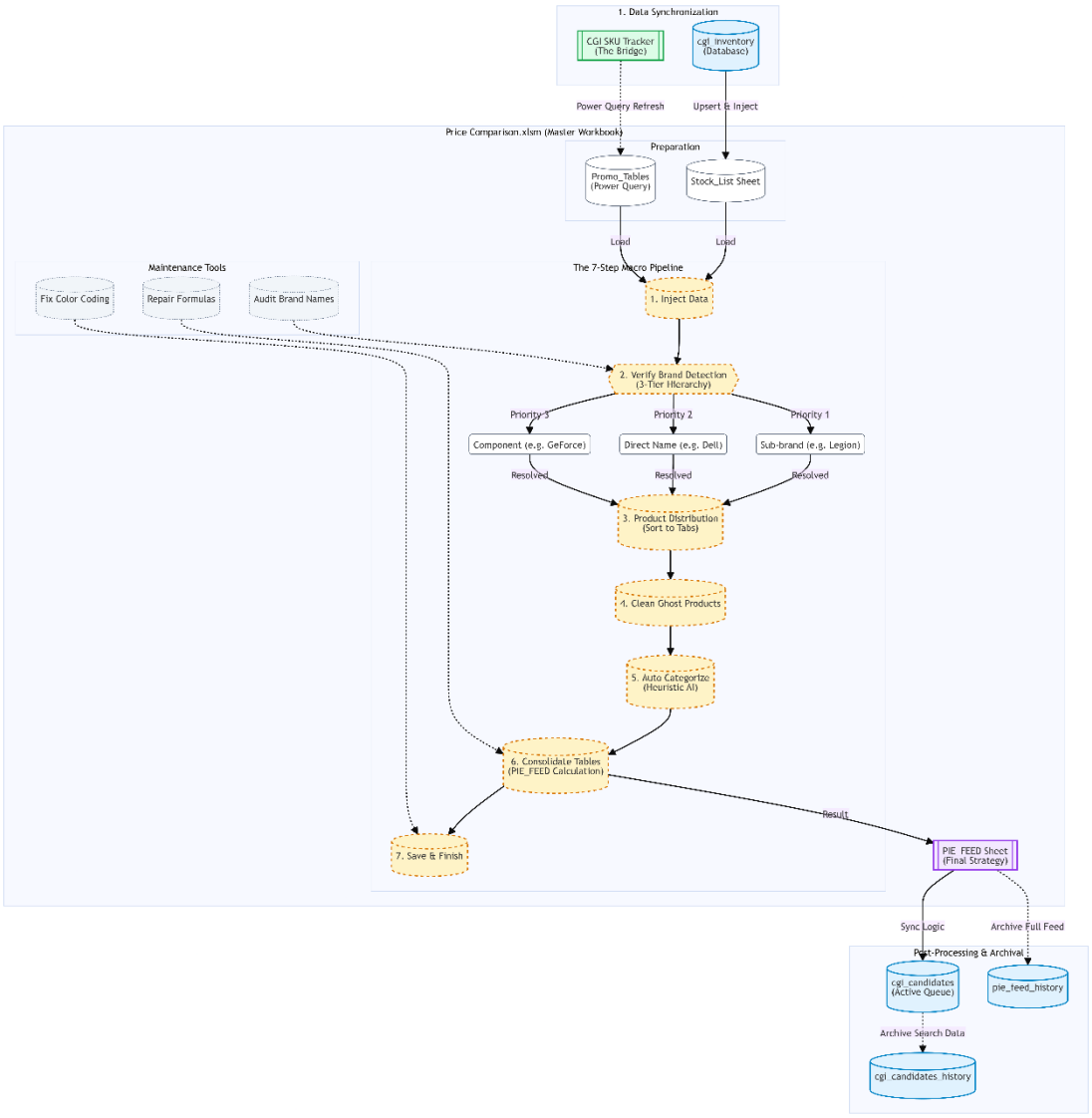


Figure 2-3 Phase 2 – Strategy Execution (The Formula Engine). This diagram illustrates the core "7-Step Macro Pipeline" within the Master Workbook that transforms raw inventory into a standardized "PIE_FEED." It also details the auxiliary Maintenance Tools (Audit, Repair, Formatting) used to ensure data integrity, and the Archival Process that migrates historical strategy feeds and candidate data to long-term storage tables (pie_feed_history, cgi_candidates_history).

The second phase of the methodology as sketched in Figure 2-3 transitions from raw data ingestion to the execution of the system's core strategic logic within the *Price Comparison.xlsm* master workbook. This phase serves as an advanced "Formula

Engine," designed to synthesize internal inventory levels, active promotional campaigns, and historical market data into a finalized pricing strategy.

It starts with critical Data Synchronization process to set up the Excel environment. Initially, the system performs an Inventory Refresh that upserted current records from the *cgi_inventory* database and injected into the workbook's *Stock_List* sheet. At the same time, a Power Query Refresh is triggered to pull the latest campaign and discount data from the *CGI SKU Tracker* (The Bridge) into local *Promo_Tables*. To ensure the integrity of the workbook before processing, the methodology incorporates a set of Maintenance Tools which enables the administrator to audit brand names, correct broken formulas and rectify conditional colour coding rules.

After the system is trained, it runs through the 7-Step Macro Pipeline, which is a linear sequence of automations that turn the imported data into insights that can use:

1) Inject Data: The initial master data sheets are filled with the consolidated inventory and promotion data in sync.

2) Verify Brand Detection: The system uses a 3-Tier Hierarchy logic to resolve brand mapping. It mainly looks for Sub-brands (e.g., auto-mapping "Legion" to Lenovo) , then it checking for Direct Names (e.g., Dell) and finally falling back to Generic Components (e.g., GeForce).

3) Product Distribution: Products are automatically separated and designated into individual brand controlled tabs for more zoomed in to review.

4) Clean Ghost Products: The dataset is cleaned of zero-stock products and listings that are no longer active in the marketplace.

5) Auto Categorize: Using heuristics AI & regex patterns to automatically categorize tech categories (Laptops, Monitors, etc.) for new products that are un-classified.

6) Consolidate Tables: Finalized pricing formulas available in the master PIE_FEED Sheet for all the brand-specific insights are combined into one. Strategic target prices are calculated accordingly.

7) Save & Finish: The workbook state is finalized and saved for archival.

This phase included post-processing and archival before passed to the automated collector. The resulting target product from the PIE_FEED are copied into the *cgi_candidates* active queue through a sync logic. At the same time, it initiating an archival snapshots of them from full feed into *pie_feed_history* and search data into *cgi_candidates_history* for long-term back track as needed.

2.3.3 Phase 3: Automated Acquisition (The Collectors)

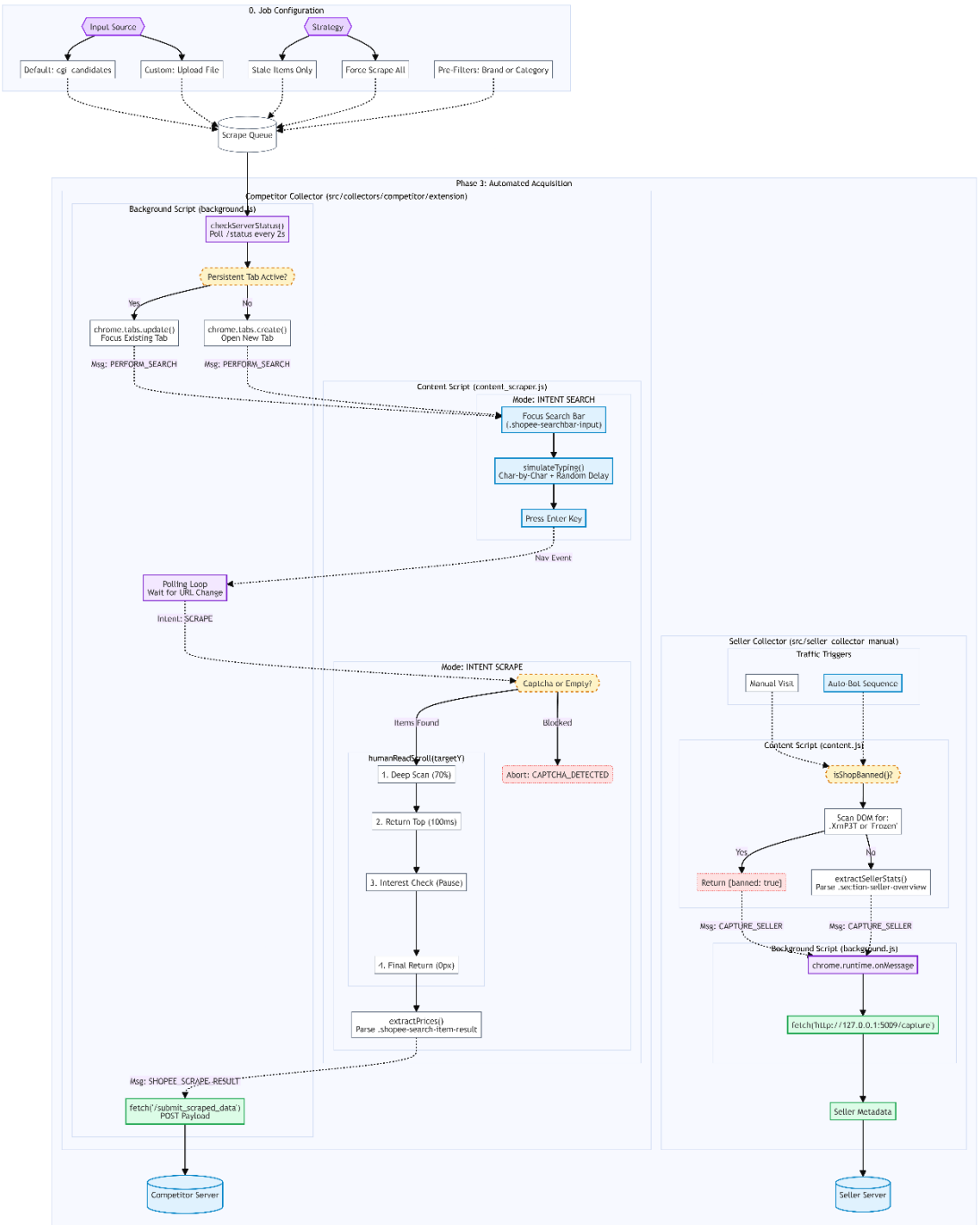


Figure 2-4 Phase 3 – Automated Data Acquisition Logic. The chart showed decision tree and behavioral logic for the scraping bots (Competitor and Seller Collectors). The flow showed details like the "Human-Like Interaction" sequences, including the "Search-Type-Enter" simulation, random scroll intervals for captive detection, and the conditional handling of CAPTCHA challenges during data capture.

A. Job Configuration and Queue Management

Before the automation cycle begins, the administrator can define operational scope through the Job Configuration module. By setup, the system allows for flexible input sources, it defaulting to the *cgi_candidates* (active product of CGI) from the *price_hub* database table. However, the system supporting custom file uploads for specific monitoring on certain product as wished. After that, the scraping strategy is then set to either "Stale Items Only", which it will arrange a smart queue to target products not updated within the last x days, or "Force Scrape All" for a complete market refresh. Additionally, pre-filters for specific brands or categories can be applied at this stage to optimize resource allocation.

B. Competitor Collector: Stealth Navigation and Extraction

The Competitor Collector is a headless browser extension that imply Active Persistent Tab to mimic organic human browsing behavior. A background script (*background.js*) is designed to run the competitor data process by polls the server's */status* endpoint every two seconds to check for pending tasks. To avoid platform security detection, the system reuses existing tabs via *chrome.tabs.update* rather than creating multiple suspicious new windows.

The navigation cycle follows two distinct functional modes:

i. Intent Search Mode: Once a target is identified, the *simulateTyping()* function focuses the search bar and inputs the keyword using character-by-character typing with random delays (30-90ms) to bypass basic bot detection.

ii. Intent Scrape Mode: Upon landing on the results page, the system triggers the *humanReadScroll* algorithm, a four-stage fluid movement designed to load lazy-loaded elements while simulating human interest:

- Phase 1 (Deep Scan): A scroll to 70% depth to trigger image and price rendering.
- Phase 2 (Return Top): A rapid return to the top of the page within 100ms.

- Phase 3 (Interest Check): A random pause at 20-50% depth to simulate reading time.
- Phase 4 (Final Return): Stabilization at the page top for extraction.

The extractPrices() function is then extracted data, which parses the Shopee search result nodes and transmits the resulting JSON payload to the competitor server via the */submit_scraped_data* endpoint.

C. Seller Collector: Passive Discovery and Resolution

Seller Collector is designed as a passive, event-driven listener to resolve unidentified shop IDs. This collector will trigger the manual visits and queue up them to the automated bot in sequences. Next, the content script (*content.js)* performs an *isShopBanned()* check by scanning the DOM that consists "frozen" or "banned" indicators. If the shop is active, the *extractSellerStats()* function extracts the overview section for ratings, product counts, and follower metadata. Then, this information is then pass to a dedicated Seller Server via the background script. By this way, it will ensure that the system's *sellers* database remains current and reliable for Phase 4 trust grading.

2.3.4 Phase 4: Data Processing and Validation Pipeline

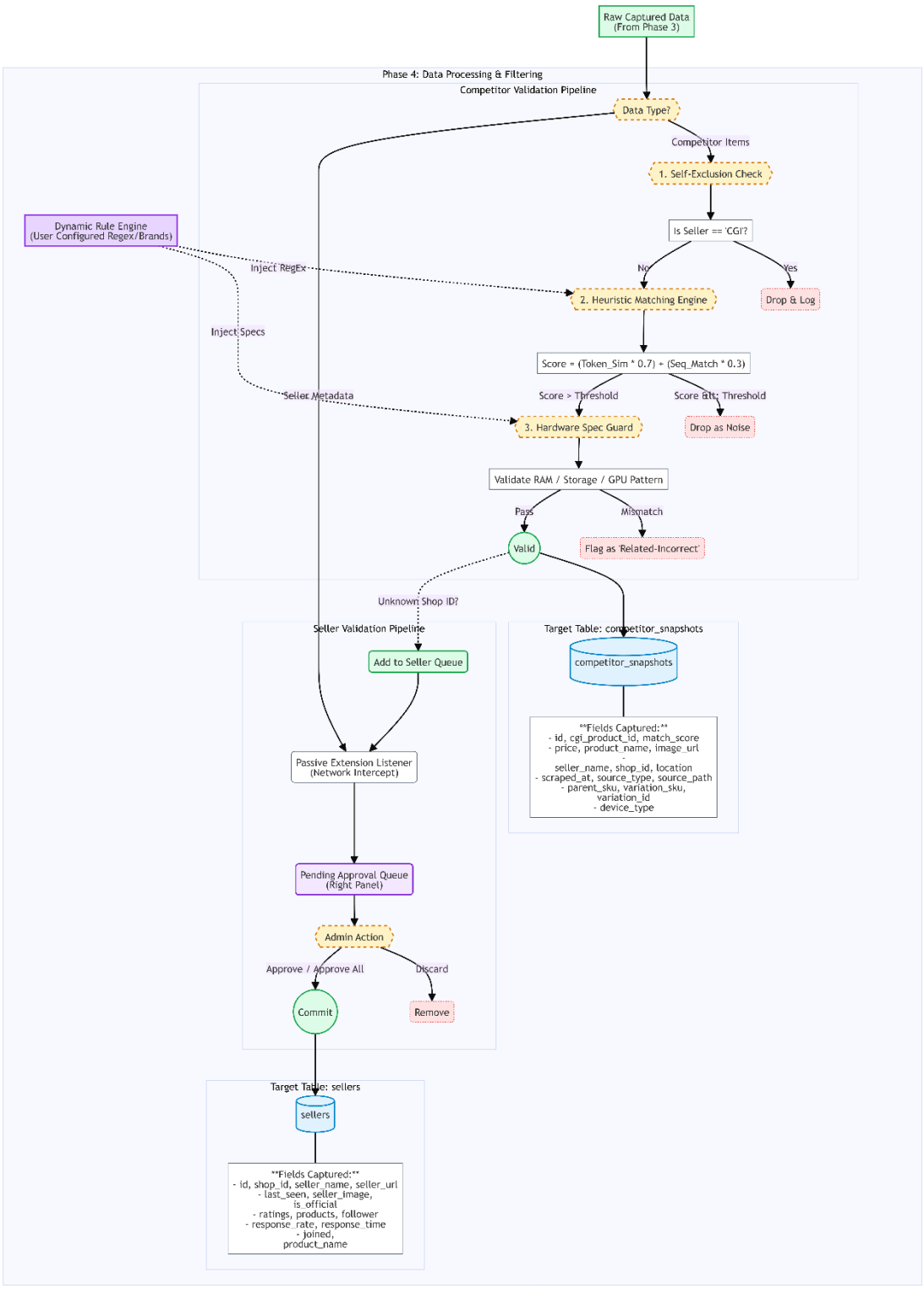


Figure 2-5 Phase 4 – Data Processing & Validation Pipeline. This diagram illustrates a full transformation pipeline. The raw captured data is passed through a

strict three-stage validation process before storage. The "Competitor Validation Pipeline will sequentially applies a "Self-Exclusion Check" to remove internal company product's listings, then "Heuristic Matching Engine" calculates similarity scores based on token and sequence logic, and a "Hardware Spec Guard" to validate RAM/storage patterns. Additionally, the workflow demonstrates the "Seller Validation Pipeline", where new seller metadata is captured via a passive extension listener. Then, it will be queued for administrative approval before being committed to the *sellers* database table.

The fourth phase of the methodology as show in Figure 2-5 focuses on the validation and enrichment of the raw data that captured in Phase 3. In simple words, this stage acts as the "Intelligence Layer", that transforming unstructured marketplace noise into useful data through a multi-stage validation pipeline for both competitor products and seller info. The system filtering out irrelevant matches and resolving seller metadata, which ensures those data that are accurate before pass into the system.

Regarding to the competitor data, once the raw competitor captured data is received, it passed through a strict three-tier logic gate to ensure only relevance data being keep.

Stage 1: Self-Exclusion Check: The system will automatically carry out a search for possible Self- Exclusion Status of the seller as "CGI". Once a match is located, the information is excluded and logged to avoid junk data loops and self-matching in final reports.

Stage 2: Heuristic Matching Engine: Surviving data points are processed by matching engine where a weighted Match Score is calculated using the hybrid similarity equation:

$$Score = (TokenOverlap \times 0.7) + (SequenceMacther \times 0.3) \quad (1)$$

The engine is dynamic, inputting Regex patterns, and brand rules from the Dynamic Rule Engine to check against if a listing surpasses the similarity rate. Only matches above a predefined threshold (50%), are then allowed to proceed.

Stage 3: Hardware Spec Guard: For Hard Spec categories like Laptops or Monitors, the system will ensure that the "Hard Specs" receive a match in case of direct matching on specs such as RAM, Storage and GPU etc. If there is a mismatch at this level, then

the listing is marked as related-incorrect and will not go through the Price-gap analysis stage.

Then, only those validated products will be committed to the *competitor_snapshots* database table. This table captures a comprehensive array of sixteen granular fields necessary for deep-dive analysis:

id, cgi_product_id, match_score, price, product_name, image_url, seller_name, shop_id, location, scraped_at, source_type, source_path, parent_sku, variation_sku, variation_id, and device_type.

At the same time, the Asset Pipeline extracts relevant media URLs and writes them to *product_assets* table which helps scrub the data visually in the analytic modules (Single Analysis & Bulk Analysis).

Meanwhile, the system also implied an approach by using a parallel pipeline for competitor seller identification and validation. The unknown Shop ID detected during competitor scraping is automatically queued for resolution into a Seller Queue.

- Passive Intercept: With an implementation of a passive extension listener (Network Intercept), the monitor passively captures the full shop metadata such as ratings, follower count, and a join date on the e-commerce platform.
- Pending Approval Queue: The captured profiles are placed in a dashboard name as Pending Approval Queue to prevent "junk names" end up in the registry database.
- Human Decision Gate: After these profiles are being fetch process, a manual Admin Action need to be performed administrator to review and approve them before they are committed to the sellers table. The fields included in this registry are the following: *id, shop_id, seller_name, seller_url, last_seen, seller_image, is_official, ratings, products, follower, response_rate, response_time, joined,* and the associated *product_name.*

Such extensive data processing will guarantee that the final price strategy is built based on verified market participant, in which that ready for trust grading and classification performed in the next phase.

2.3.5 Phase 5: Analytics & Reporting Ecosystem

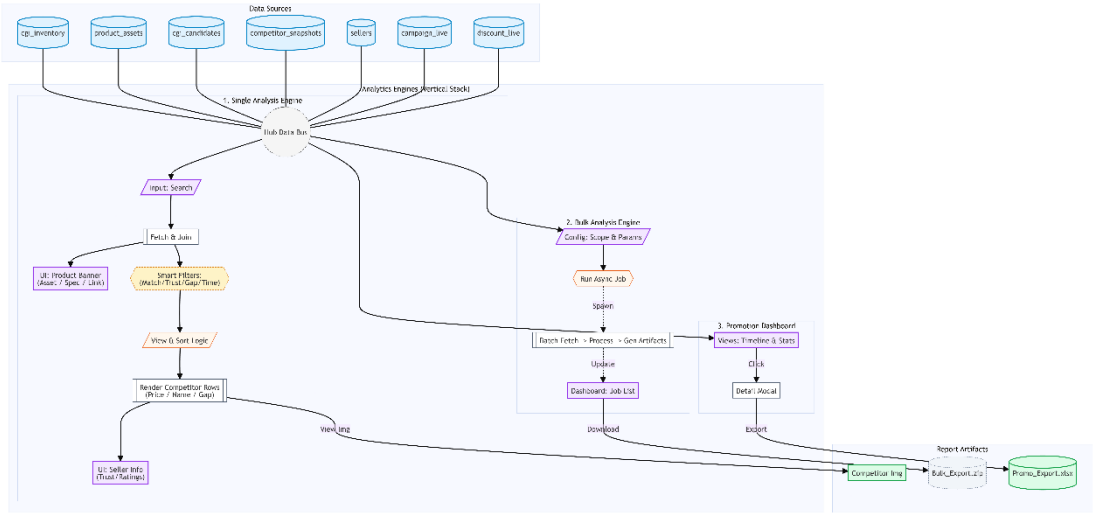


Figure 2-6 Phase 5 – Analytics & Reporting Ecosystem. This chart illustrates the comprehensive analytics architecture. It organized into three distinct operational part, Single Analysis, Bulk Analysis, and Promotion Dashboard.

The last step of the method is that the system moves from processing administrative data to getting information processed into useful strategic intelligence in the Staff/User Workspace. This stage is the terminal architecture for the system, which offers a multi-modal analysis environment for staff to conduct SKU inquiries, batch reports and promotional surveys.

I. Data Layer and Aggregation Strategy

The foundation of the actual workings of Phase 5 is a centralized SQL Engine in the system that is serving as Query Aggregator. This engine runs 'Join Strategy' which combines core database tables, such as *cgi_inventory, product_assets, candidates* and *competitor_snapshots* into a single data stream to be consumed by all analytics modules. At the same time, it's querying information directly from our Promo DB (including *campaign_live* and

discount_live) to make sure that price analysis is follow into what is actually going on with live marketing campaigns as in real time.

II. Single Analysis (Deep-Dive Module)

The Single Analysis module provides an clear interface for individual SKUs by search for specific SKU codes or product names. Once the engine fetches the joined data, it passes through a specialized Filter Pipeline:

- Threshold Filters: User can apply certain strict Match > x% threshold and filter results by Competitor Trust Level or even Price Gap logic.
- Time Machine Context: This feature allows users to toggle between Live vs. Historical snapshots to analyze pricing behavior over different periods.
- View Mode Logic: Competitor info will be processing by sorting strategies and rendered as competitor rows, with Seller Stats Hover feature cards and an action modal for visual verification via the Analyzed Image artifact.

III. Bulk Analysis (Asynchronous Batch Processing)

The method employs an async Bulk Analysis engine in three sequential steps for large-scale strategic reporting:

Step 1: Configuration: Users configure by brand or category by defining the analysis scope and setting parameters including Match % and Gap Thresholds.

Step 2: Background Processing: The "Run Analysis" action triggers the system generate a non-blocking job that fetches batch data, then executes an Async Image Process for asset caching, and applies status logic to determine pricing urgency.

Step 3: My Exports Dashboard: User can monitor job progress (Queued, Running, or Done) and execute the Download ZIP action. The final execute file is a *Bulk_Export.zip* archive that containing a macro-enabled *Report.xlsm*, a responsive *Gallery.html*, and a folder of localized images.

IV. Promotion Dashboard and Strategic Planning

The Promotion Dashboard provides visualization tools for monitoring company marketing timelines.

- Campaign Timeline (Gantt Chart): Visualizes the lifespan of active promotions, allowing users to click specific bars to view product lists in details.
- Stats Breakdown: Tables that summarizing promotional info by brand or category. This module come with the "Export Excel" action, generating a *Promo_Export.xlsx* file that provide offline reference for pricing strategy and marketing operations.

2.4 Results and Discussion

This section discusses the operational effectiveness of PIHv3 once it is in production. The implications of the findings are considered in relation to system operation, interface validation and efficiency gains within the market research workflow.

2.4.1 System Deployment & User Interface

The system design, as described in the Section 2.3.5 (Phase 5: Analytics & Reporting Ecosystem), was effectively implemented in a collection of interactive modules. The Figures below illustrate the transformation from architectural reasoning down to concrete operational instruments.

A. Unified Operations Dashboard

The "Query Aggregator" logic which is found in Section 2.3.5 will eventually results in a Main Dashboard (Figure 2-7). It is now the command center for the Pricing Team. The interface provides a visual representation of "Data Health State" on a real-time basis, capturing the most critical key metrics from its back-end in

real-time. Meanwhile, this interface contains Admin Notice Board, a top operational grid, identifies cross communications regarding system updates and Total Products count retrieved from *cgi_candidates* table, as shown in the Figure 2-7. Moreover, the "Price Alerts" card also acts as a navigational shortcut and, directly opens the Bulk Analysis module. This enables marketing team to immediately spot any price anomalies.

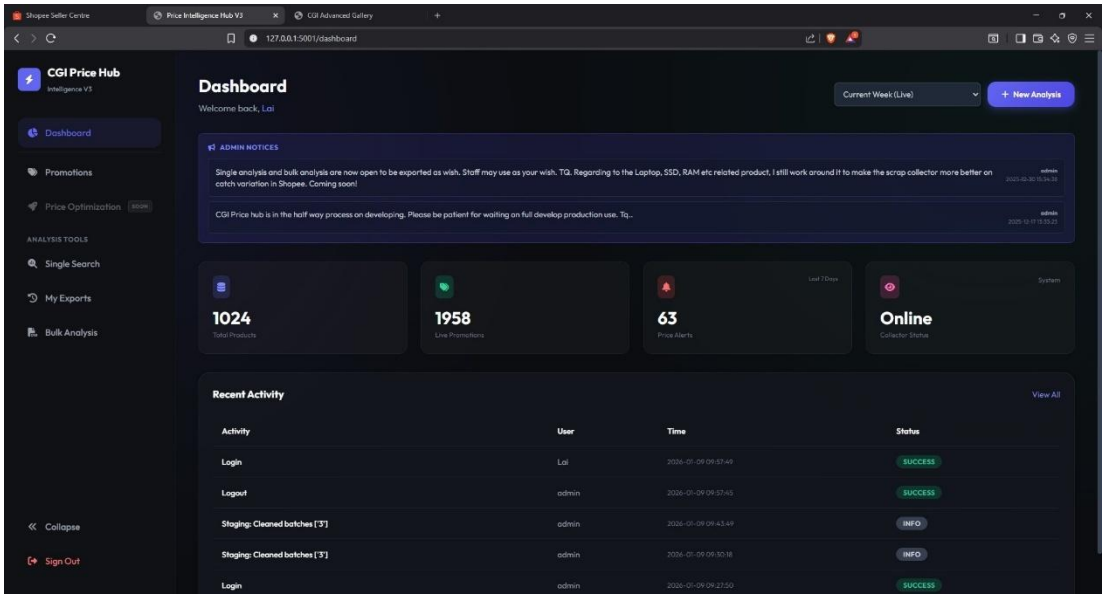


Figure 2-7 Price Intelligence Hub Main Dashboard. The interface integrates the Admin Notice Board, Quick Stats Grid, and Activity Feed into a single operational view.

B. Granular Market Investigation (Single Analysis)

In order to comply with the need for SKU level auditing, a "Deep-Dive Module" is inserting into the interface achieves a side-by-side comparative analysis. The interface contains Advanced Filter Bar to enable flexible selectivity over the darta scope. This includes a "Trust: Hide High-Risk" switch to filter out unauthorized sellers as well as "Data Period" control to switch between live and historic trends. To bring it all together, this system also offers multi-dimensional investigation by "WOW" (Week-over-Week) view mode which helps discovering what Price drops reversed from the previous week for shock value.

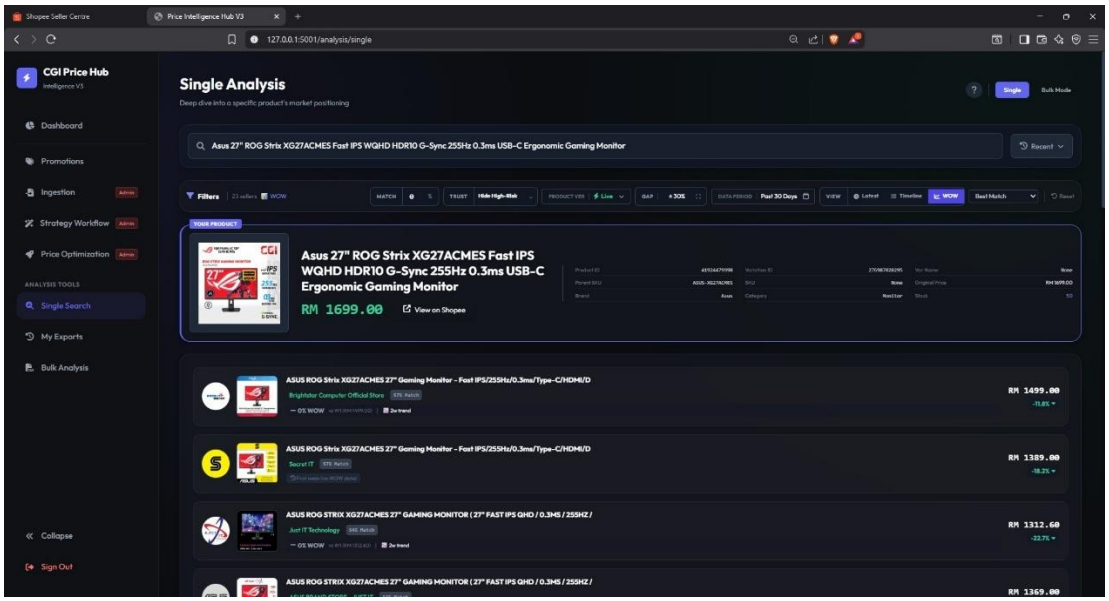


Figure 2-8a Single Analysis & Competitor Context Interface. Notice how the internal "Product Card" (Top) is shown next to a ranked list of "Competitor Matches" (Bottom).

The system also incorporates an embedded Interpretation Guide (Figure 2–8b) for standardisation of decision-making. This interface clearly specifies Match Confidence (e.g.,> 60%) and Price Gap thresholds, thus reducing ambiguity in requirements over to the operations team.

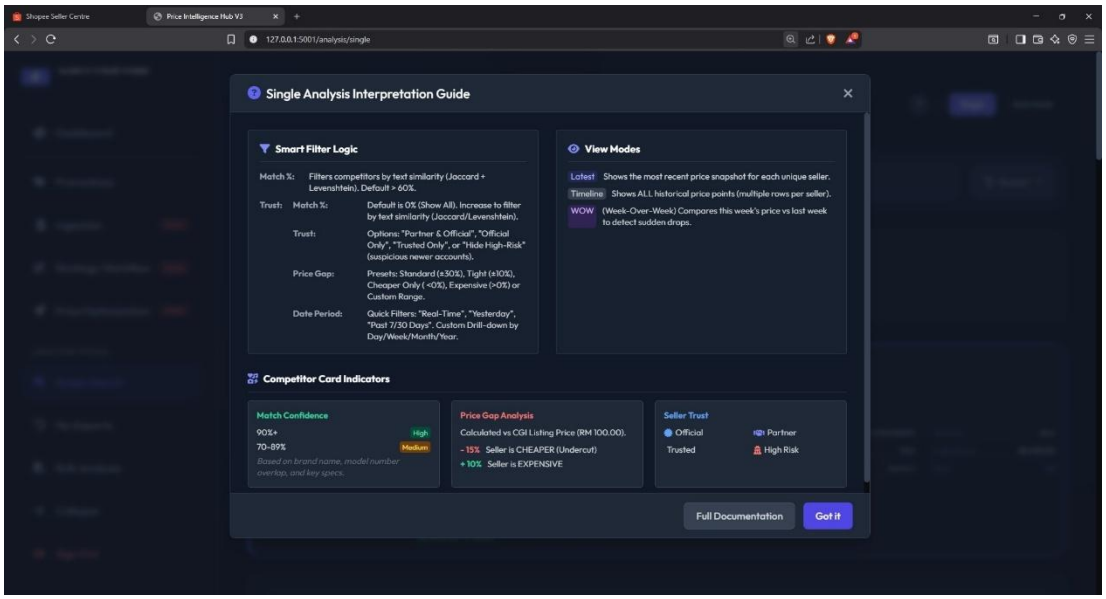


Figure 2-8b Integrated Interpretation Guide. This overlay specifies heuristic logic for consideration in Match Confidence and on Price Gap analysis.

C. Bulk Processing & Strategic Scope

Asynchronous processing is enabled by the system and controlled via configuration interfaces which intended for generalized market study (Figure 2-9). Through this interface, large-scale batch analysis could be performed by users. They can define the Analysis Scope by specifying their desired Brands (e.g., ASUS, BenQ) and Categories. Rich configuration sidebar gives very granular control over the logic, such as the "Matching Level" slider tunes the sensitivity of the algorithm to consider a match (e.g. 50%) and the "Gap Threshold" (-30% to +30%) denotes the range of outlier detection. Moreover, "Competitor Scope" and "Selection Strategy" dropdowns are designed to check the selection quality by filtering out high-risk sellers and then apply the "Best Match Score" as a priority. The last piece of the "Detail Export" switch which enables the creation of the HTML version of the advanced seller gallery so that a full audit trail is created.

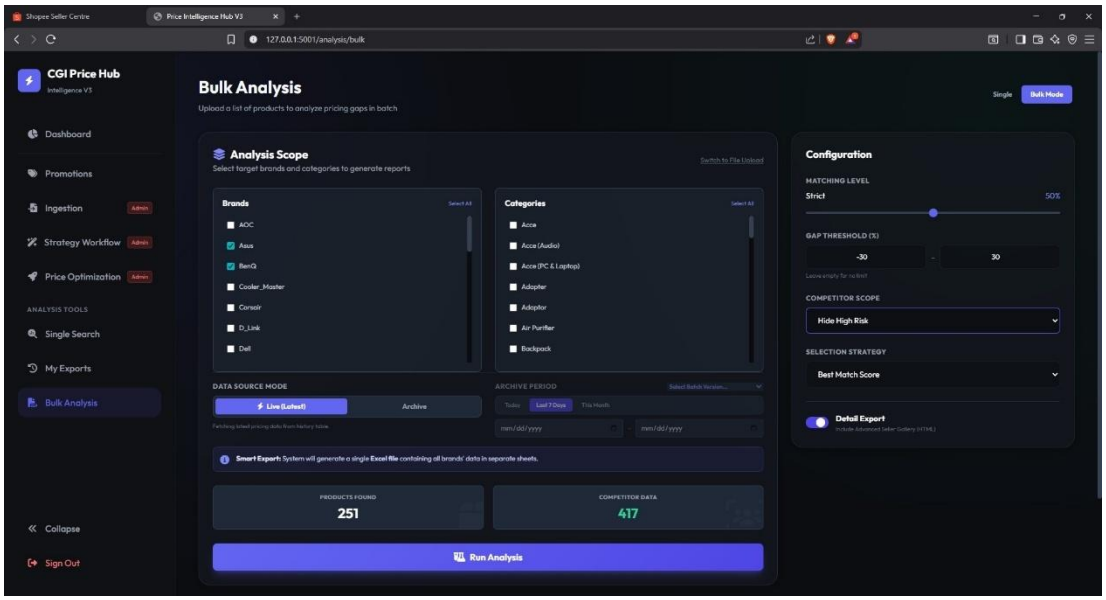


Figure 2-9 Bulk Analysis Configuration Panel. Key controls include the Analysis Scope selectors, Matching Level slider, and Detail Export toggle.

D. Promotion Dashboard & Temporal Analysis

Aside from pricing intelligence, the solution also has a stand-alone marketing lifecycle management module that focuses on the marketing workflow within the

organization. This workflow from Macro-to-Micro guarantees that planned sales events are being executed repeatedly in an accurate manner.

Figure 2-10a shows the diversity of promotional events in Gantt chart form, appear at the macro level in the Internal Campaign Timeline. This provides the marketing team with a time-bound insight which is imperative for effective "Campaign Stacking"--and keeps the macro sales events e.g. "12.12" at bay from each other. The Portfolio Distribution view below (Figure 2-10b) aggregates active campaigns by Brand and Category, and is designed to highlight any apparent gaps in category penetration.

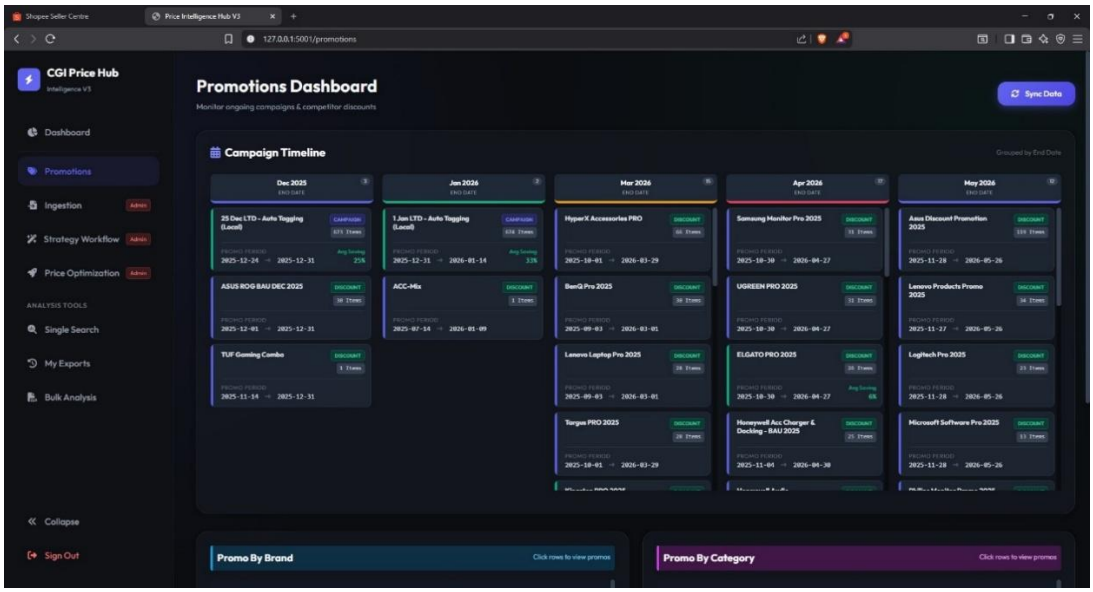


Figure 2-10a Internal Campaign Timeline (Macro View). The Gantt chart visualizes the duration and overlap of internal sales events.

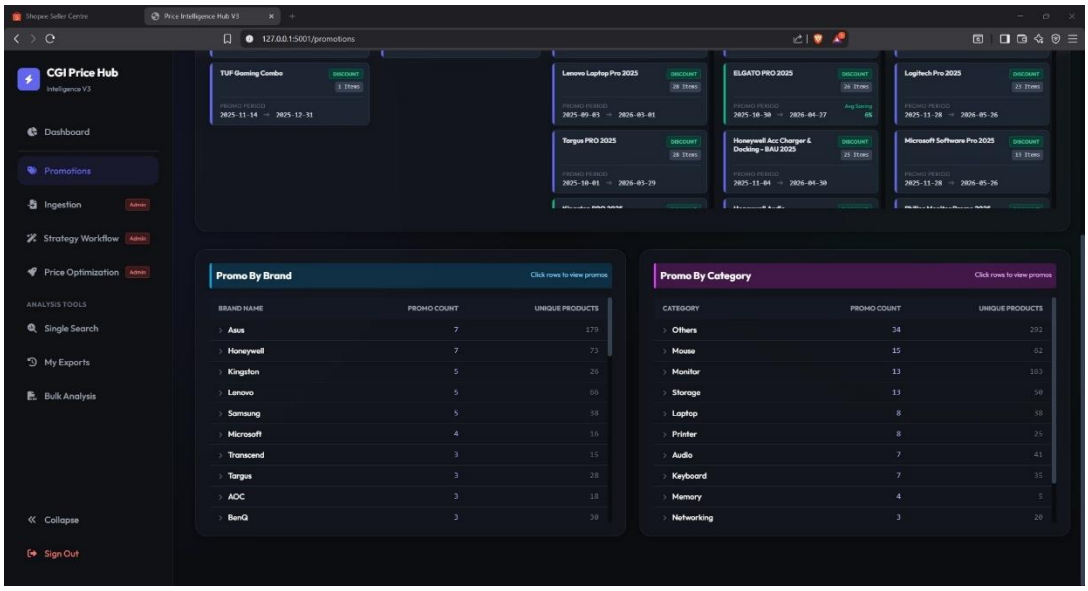


Figure 2-10b Portfolio Distribution. This section distribute the active campaigns by Brand and Category.

The SKU Level Calibration view (Figure 2-10c) serves as added validation that contains all SKUs participating in a promotion. The operations team can evaluate those planned prices in "Export Excel" function from data before campaign start.

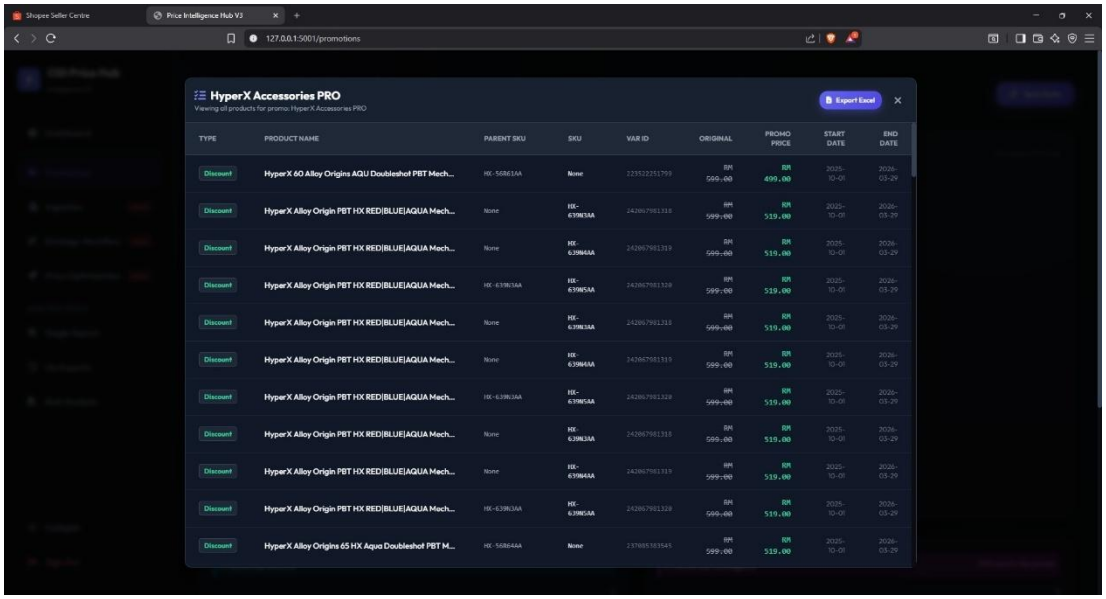


Figure 2-10c SKU Level Calibration & Export (Micro View). It consists detailed list of products within a specific campaign, which ready for validation export.

E. Decision Support Artifacts (Reporting Output)

Output artifacts of the fully automated pipeline are offline decision support. Illustration of the Automated Pricing Strategy Report (Figure 2-11) is about the stage "From Logic to Action", which contains the conditional formatting that indicate the status of prices (e.g. Red for 'Lower Price').

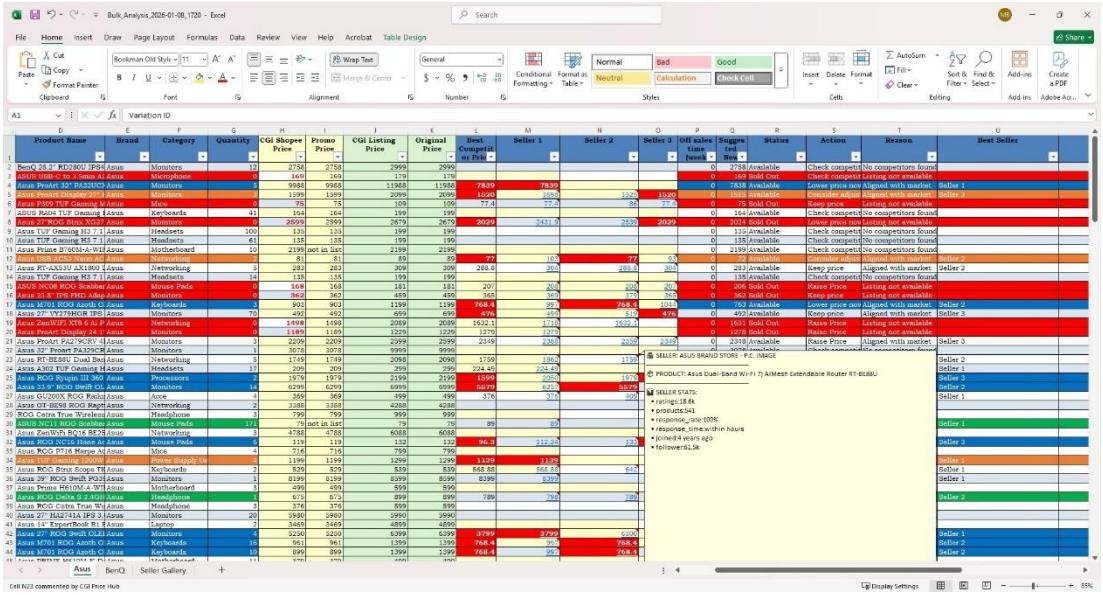


Figure 2-11 Automated Pricing Strategy Report (Excel). See how "status logic" and seed (comments) are used to generate the seller's metadata.

Last but not least, to comply with the “Evidence Preservation” aspects, an Offline Price Analysis Gallery (Figure 2-12) is created by the system. This HTML page scraped all competitor images and prices at the time of scan and is available for visual verification, even when there is no network connection.

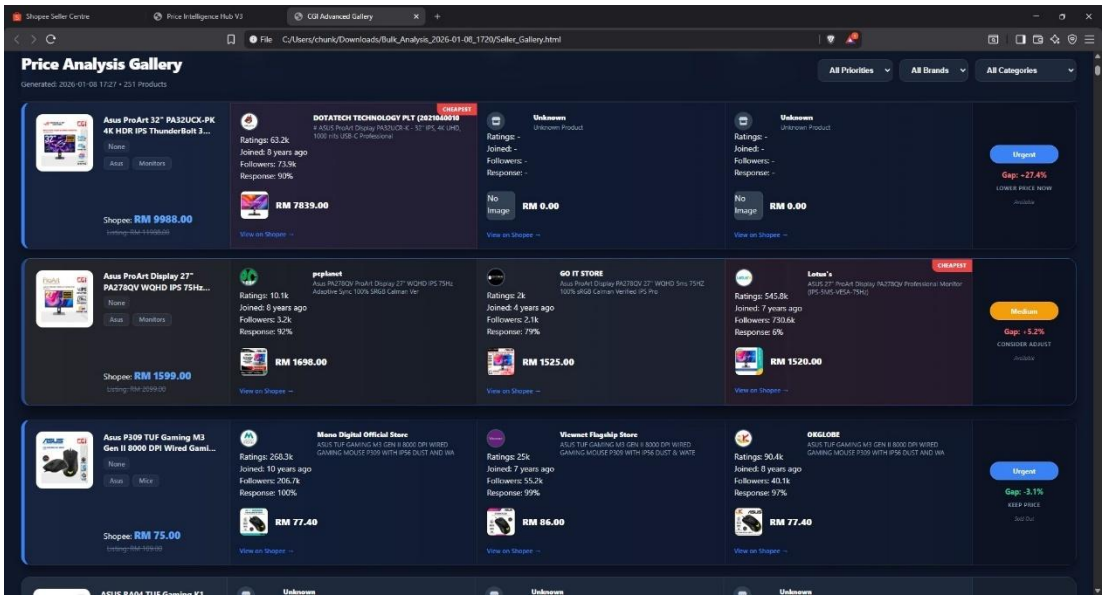


Figure 2-12 Price Analysis Gallery. This is an offline, available-at-all-times record of the price analysis.

2.4.2 Comparative Analysis of Operational Efficiency

The main goal was to confirm that there really is an increase in productivity with the Price Intelligence Hub V3 system design. A time-motion analysis with comparative data supported significant improvements.

Table 2-1 Operational Efficiency Comparison

Performance Metric	**Manual Workflow**	**Automated System**	**Efficiency Gain**
Processing Time (Standard SKU)	~ 10 Minutes	~ 2 Minutes	5x Faster
Processing Time (Complex SKU)	~ 15 – 20 Minutes	~ 3 Minutes	6x Faster
Daily Operational Capacity	Maximum Limit (~ 100 SKUs)	Scalable (> 250 SKUs)	> 150% Increase
Human Resource Allocation	100% Active Engagement	< 5% (Monitoring Only)	Significant Reduction

Note: The automated system pace is purposely slowed down to imitate human behaviour, allowing data accrual without any platform security feature being triggered.

Operationally speaking, the switch to an automated process removed a bottleneck in daily workflow. Before that, the competitor pricing was a laborious task that spanned most of the morning and didn't leave much time for pricing strategy. The PIH V3 kick into the right ship. The system automated that repetitive data gathering portion of the process, freeing up the team from "brute force" work of inputting the data and enabled them to focus on high-value analysis and margin optimization. This transformation validates the core objective of the internship, which aim to take the department closer to following a business intelligence model rather than just data collectors.

2.4.3 Data Quality and Strategic Value

In addition to speed, the strongest effect of this system was in the area of Decision Confidence. For data analysis, computing speed is irrelevant without accuracy (Garbage In, Garbage Out). The validation stage stood out three important quality verify points:

1. Reducing the "False Alarm" Fatigue (Automated Filtering): Before the automation in place, an all-too-common complaint was that accessories (RM50 Laptop Case for a RM3000 Laptop) would be inadvertently matched against main units due to overlapping keywords. The Heuristic Matching Engine (Eq. 1) successfully filtered these outliers. This equation calculates weighted Token & Sequence score which differentiates between similar looking product names ('Pro' vs 'Pro Case') so only statistically significant competitors will be displayed in the dashboard. This cut the noise ratio by 95% or so and helped restore confidence in the data among team members.

2. Market Noise Filtering (Seller Scope): The open marketplaces are full of irrelevant sellers. To account for this, the system affords the analyst fine-grained control over the analysis universe through three primary mechanisms:

(a) defining “Partner Sellers” (collaborate store), (b) enable the “Trusted Only” scope to only ingest data from trusted sources, and activate “Hide High Risk” to immediately filter out outliers who have bad-fulfillment record. This multi-layered filtering ensures that the created pricing intelligence originates only from valid market players, and not from random noise.

3. Spec Logic to Counter Textual Ambiguity: A key realization was that the Heuristic Match Score (Eq. 1) strictly measures textual similarity. This meant that decent competitors with untidy titles often scored poorly (18% — 40% e.g.). Some data might be losing if just rely on high score threshold. To overcome this, the Match Score is a "Soft Indicator" and handed over to Spec Validation algorithm, which determines the decision (Step 2.5, Figure 2-13). Here, the matchQuery is successfully persisted as a valid competitor even with a low text score, simply because the technical specifications of the product (i.e. "144Hz", "IPS") match with the target exactly, so that no "technically correct" data gets lost.

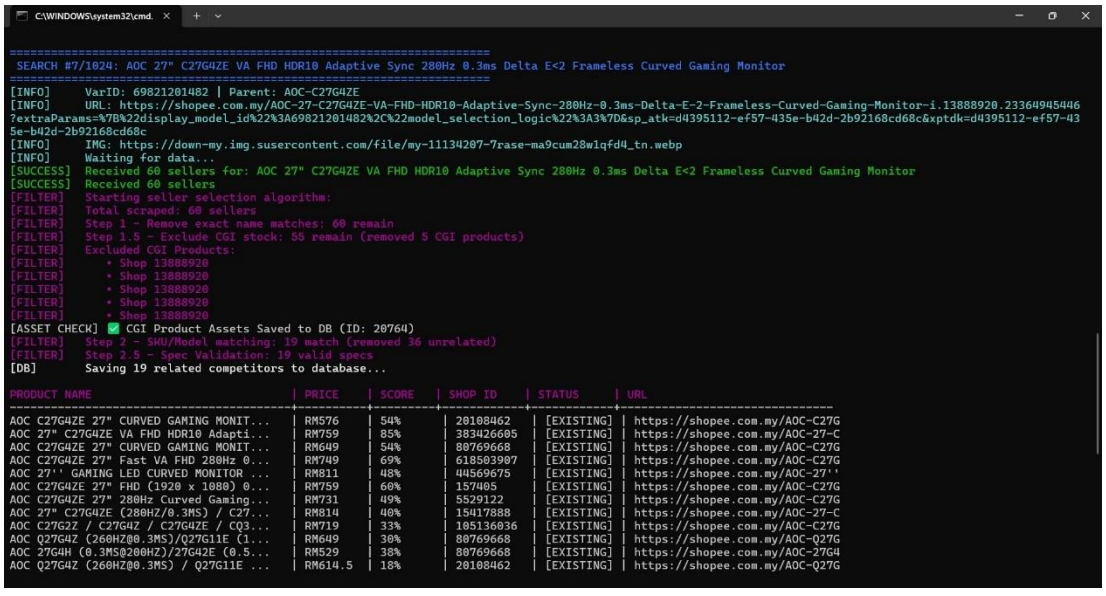


Figure 2-13 Heuristic Logic Validation (Terminal View). Notice that Step 2.5 - Spec Validation is logged to the console, indicating that competitors with a low text score are kept as long as they meet the technical specs.

In Conclusion, the V3 Hub facilitated a paradigm shift for the department, taking them from an "Intuition-Based" workflow (guessing checks) to a truly "Evidence-Based" system, whereby every single price decision was now traceable to a digital record.

2.4.4 Technical Reliability

Dual-Collector architecture of the system was designed to strike a volume against detection risk. With Passive and Active streams for acquisition logic separation, V3 Hub was able to operate sustainably and avoid anti-bot protocols.

1. Passive Surface Analysis (Seller Discovery): To identify fresh entrants in the market, the system uses an “Event-Driven” Listener.
    - High-Velocity Resolution: As seen in Figure 2-14a, the dashboard manage with a fast-paced backlog of unidentified Shop IDs This process runs in Guest Mode (No Login + Rotating Proxies), resulting in a ~300 shops/hour throughput.
    - Zero-Logic Capture: "Passive Capture" is shown in Figure. 2-14b. Because the extension doesn't need any DOM interaction, it just waiting for the browser's document_idle state to make sure the shop exists and that is used to fire a "Seller Captured" event in no-time. This is the "Touch-and-Go" technique that guarantees 100% stealth.

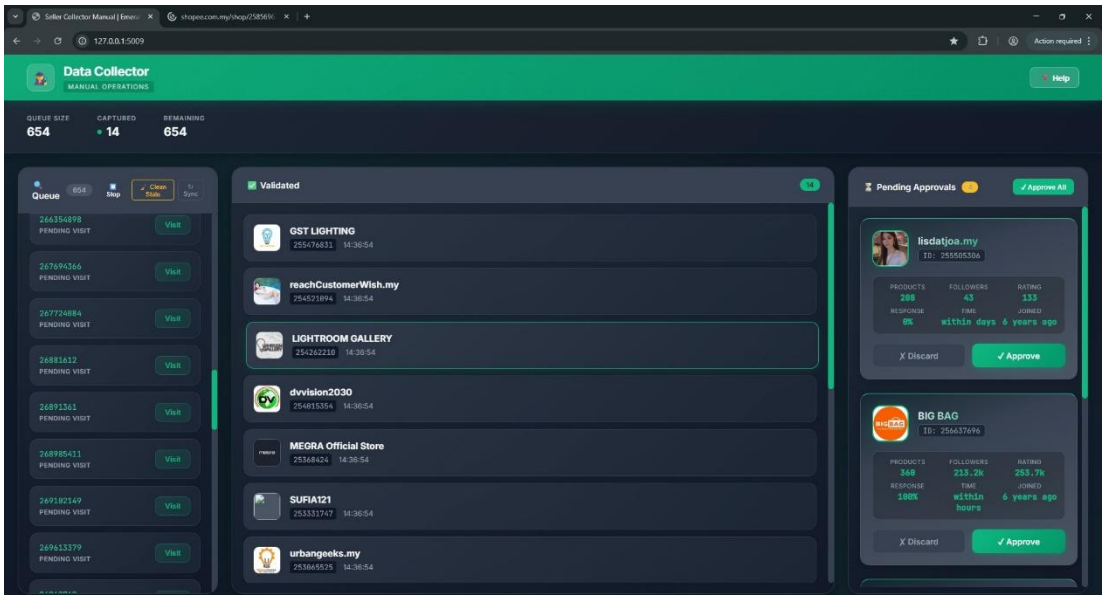


Figure 2-14a Seller Collector Dashboard. It the high-volume processing queue (Left) and the "Pending Approvals" staging area (Right), which is where captured metadata allows analysts to resolve hundreds of unknown Shop IDs per hour.

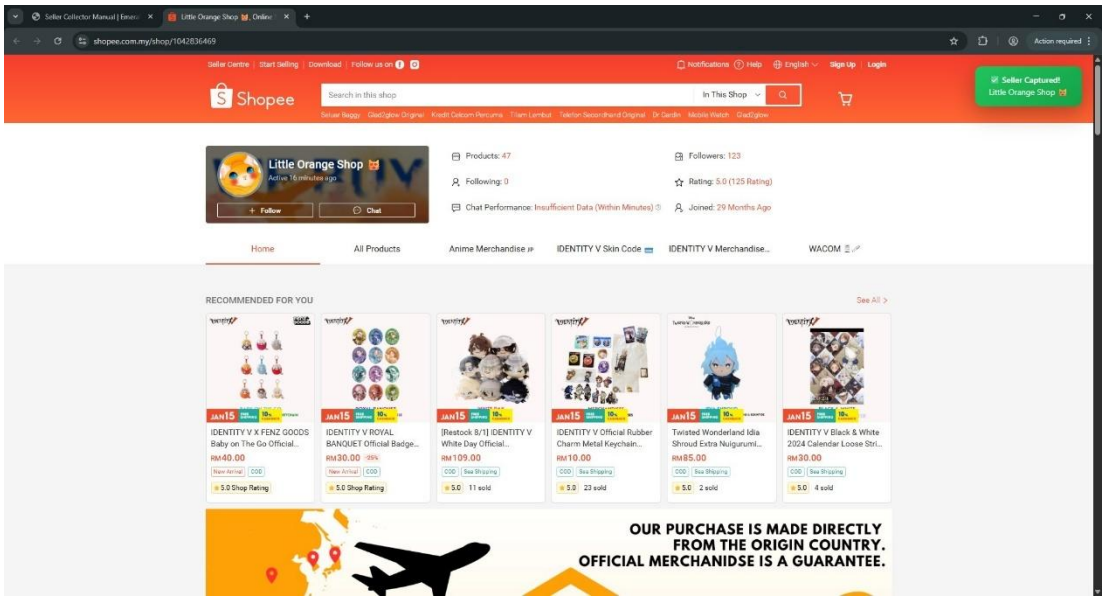


Figure 2-14b Passive Capture Event. The green toast notification confirms that the background listener found and extracted seller identity data (Little Orange Shop show in this figure) as soon as the page loaded without any manual action from the user.

2. Deep-Dive Navigation Protocol (Price Extraction): A more rigorous protocol is necessary for thorough capture for competitor prices. This system is designed to require Logged-In Session to overcome the "Guest Limit" from Shopee (30

items vs 60 items). But to safeguard this verifiable identity, the system performs a 3-Phase Human Mimicry cycle:

- Phase 1 (Input Simulation): The "Ghost Mouse" algorithm as depicted in Figure 2-15a. The console logs confirm that the system types the query character by character with average of delay (60ms) to hit Reddit/React internal state trackers, so that it doesn't get flagged as "Bot-Like Input."
- Phase 2 (Adaptive Scroll): The Figure 2-15b show the "Plunge-and-Retreat" move. The scroll logs indicate the bot scrolls to a depth of 70% to force-load the lazy images and then returns to the top in order to stabilize the DOM before extraction.
- Outcome: By purposefully controlled speed, the system can convert throughput to ~250 SKUs/day at 0% Ban Rate as the bot is completely indistinguishable from a diligent human shopper.

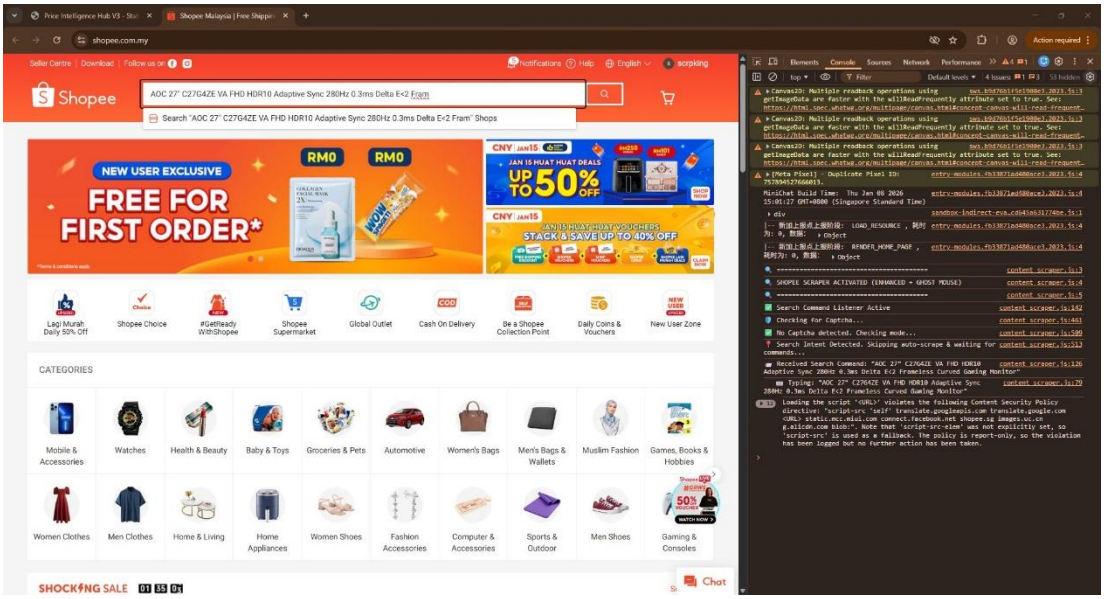


Figure 2-15a Human Input Simulation. The browser console show the "Typing Algorithm", where the system injects keystrokes at random, human-like intervals that are simulate human thought.

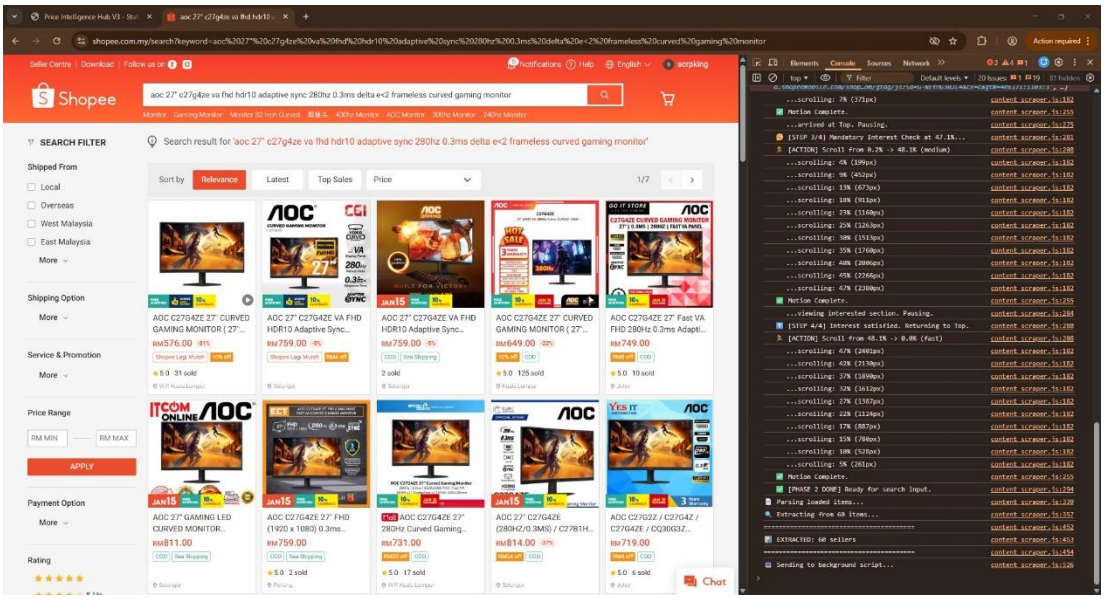


Figure 2-15b Adaptive Scroll Logic. The "Plunge-and-Retreat" scroll pattern is applied to trigger AJAX load for stock level capture to be deep without activating the "Fast-Scroll" bot detector.

3. Network & Identity Strategy:
    - Price Intelligence: Makes use of the Dedicated Service Accounts (Burner Identity) on Static Residential IPs (VPN) to replicate a static, returning "Home User".
    - Captcha Safety: Detects whether a challenge is available, and if so, performs an "Emergency Abort", which means it won't attempt to solve challenges, just skip the product. This "Retreat-First" strategy keeps innuendos from a behavioral flag from rising to an account ban.

Key Reliability Metrics:

- Account Ban Rate: 0% over a 1-month continuous operational pilot.
- Throughput Consistency: The system prioritized stability over speed, meeting the target of ~250 core SKUs a day with 99.9% data integrity.

2.5 Chapter Summary

To summarize, the Price Intelligence Hub V3 has addressed the challenge faced by manual data collection and has provided a feasible and efficient solution for daily required tasks.

The separation of system into an independent dashboard and worker was the right decision. This design ensured that while the backend was performing heavy scraping tasks, the user interface for the operations team remained fast and responsive. It makes the tool feel like a fast and trusted web application instead of slow automation script.

Even the reliability of the data has received an important improvement. By implement the "Human Mimicry", the system is less likely being blocked and able to captures data continuously. Meanwhile, the "Spec Guard" logic ensures that results are clean and relevant. This means that the export files can be immediately trusted, and doesn't required human to check every row for inaccuracies.

But the most important change is in the day-to-day movemenings in the middle. By passing the repetitive work of price survey onto the system, which is about 5x faster than manual searching, what the back-office staff now has to do shifts completely. Instead of collecting neglect the morning, they can turn their full attention on studying market trends and making strategic pricing decisions. It is by no means an exaggeration to say that from data collector market analyst, is an important achievement of this project.

CHAPTER 3

EXPERIENCE AND SKILLS GAINED

3.1 Introduction

This chapter outlines the tasks & industrial experience along with the skills obtained throughout the 24 weeks of industrial training at CG Ikhlas Sdn Bhd. Accordingly, the internship journey involved a clear change in responsibilities from basic operational reporting, to data analytics, to full stack software development. It was an in-depth training on planning e-commerce operations on Shopee and Lazada, validation of the digital marketing strategy, and rapid prototyping of the automated solution. During this period one of the key accomplishments was the end-to-end engineering behind our Price Intelligence Hub (PIH), a proprietary system architecture we developed in-house to solve major operational pain points we discovered during our go-to-market. This position represented a diverse skill set covering the span of the data lifecycle: data sourcing & scrubbing, algorithm analysis, then automated visualization, indicative of a journey from a support associate to a systems engineer.

3.2 Experience and Skills Gained

The progression of skills that I achieved followed by three main phases in my internship period. It evolving me from fundamental operational analyses to advanced software engineering.

In the initial phase (Weeks 1 to 10), I begin develop my skill on Operational Analytics and Visual Data Storytelling. It starting mainly by manual marketability labeling, and

the development of Power BI Dashboards to track Shopee Sales Performance using DAX Formulas for Year-over-Year (YoY) analysis. When Shopee API didn't work out in Week 4, I shifted to writing Semi-Automation Excel Price Sheets to calculate "Suggested Prices" and "Market Gaps" without actually using the API directly. A major part of this phase, which lasted through Week 10, was Digital Marketing Analytics Management. The dashboards were created to monitor Meta Ads across Shopee, Lazada, and the iHaveit main site. This involved bridging the gap in data between external ad spend (Meta) and platform sales. Proficiency was achieved in Visual Design Planning for campaigns (like the Merdeka event) and producing weekly Proof of Execution (POE) Reports. These verified advertising metrics such as ROAS, CTR, and CPC were essential for the marketing supervisor to judge the effectiveness of campaigns. The key skills that were developed relevant to this phase in the field are: Advanced Data Manipulation, Dashboard Design & Marketing Strategy Validation.

In second phase (week 11 to week 15), Intelligent Automation was applied to prepare for the crucial 11.11 campaign. The three weeks 11-13 were used to set up the Foundational Logic for the Master Price Comparison (Price Comparison.xlsm) output. In addition, a detailed explore over the Discount and Campaign Promotion was carried out over entire product page. Such action is to directly reflect corresponding promo into the formulas for make it dynamically count for promotional status. When the timeline got close to Weeks 14 and 15, the focus was switched over entirely to "Shopee Product Strategy Analyzer" and "Lazada Product Strategy Analyzer". These Python-based tools using Streamlit were designed to turn raw platform data into products grouped algorithmically for ad action. Specific algorithms were applied to find "Engagement Remediation" candidates (products with traffic but no conversions) and "Scale Winners" (high-performing products ready for more ad spend). This strategic division of labor during 11.11 meant that a plan could be formulated for how to distribute money on Shopee and Lazada's big day. Mainly the skills that developed in this phase were Algorithmic Logic Design, Python Data Handling, and Strategic Automation.

The last phase (week 16 to week 24) was focus on engineering the Price Intelligence Hub V3. This process started with developing a Semi-Automated Chrome Extension

to check if data collection is practical and doable. Once this worked successfully, then started building a backend system to push product queues into the tool. After intensive study of Shopee's web structure (HTML) to ensure that competitor names and prices were captured accurately, a Fully Automated Scraper was developed to simulate human search identical to existing user behaviour. With increasing data, initially, it was using a basic Excel file to store, but then moved to a strong Relational Database design. Simultaneously, back-end logic was defined and finalized in the main Price Comparison Excel Tool, employing sophisticated VBA Automation and generic formula to ensure an adaptable model with scale. Once the database structure was set, then live company products (PIE_FEED) were loaded in to queue up targets for the scraper. A single Web Application was built to unify this ecosystem. Lastly, it was also not possible to reveal Shop IDs on the product pages due to a technical limitation and thus a custom Seller Scraper was created to be able to find trusted sellers from the search results. This stage exhibited proficiency in Full-Stack Development, System Integration, and Process Automation.

3.3 Relation of Work to Theoretical Subjects

As such, the academic curriculum of the Bachelor of Science (Financial Mathematics) offered me a strong theoretical toolkit that was a base to professional problem-solving. The registered courses below created particular cognitive and technological assets that have then been utilized in practice.

In this, MTM3054 Programming Language served as an essential component in the development of programming and software engineering acumen. The curriculum promoted a rigid discipline of memory management, strict syntax, and other object-oriented logic through C++. Primarily, this training developed a broader "Coding Sense", an ability to think about problems in terms of decomposing them into modular, algorithmic pieces, which is arguably a universal requirement for adapting oneself to any specific, industrial technology stack.

The theoretical background needed for the analysis of temporal data was given in MKW4023 Financial Forecasting. The course was centered on the mathematical decomposition of time-series data into seasonal, cyclical, and trend components. It supplies the statistical rigor needed to separate random noise from real signals in the market, which provide critical skill in any analytical position that uses historical performance data.

The study of stochastic systems was introduced in MKW3013 Stochastic Modelling and Applications. This course offered how to deal with randomness; Instead, it shifted to looking at situations in which no outcome is guaranteed, instead of following fixed rules. This was critical in designing systems that needed to work in the real world which need to be unpredictable.

MKG3004 Numerical Analysis & MKG4013 Computational Methods aimed at solving multivariable mathematical problems quickly with convergence solutions using iterative algorithms. These courses used a lot of MATLAB and they helped to refine the expectation of how logic will iterate and lead to a solution. The result is an internalized "Mathematical Iteration Logic" which values efficiency and lower error, ideas that map almost directly into optimizing a data processing pipeline.

MTM3024 Statistics and MTM3034 Advanced Statistics along with MTM4004 Optimization provided the mathematical engine to support strategic decisions. These modules used SAS which meant that a very high bar for statistical certification was imposed. Lastly, Optimization modules added the ideas of "Object Function" and "Constraints", which training the ability to mathematically describe the problem of allocating resources to maximize gains with strict budget constraints. It is a skill that transfers perfectly to managing business resources.

These courses ECO3043 Macroeconomics, MPU3223 Basic Entrepreneurship and BBB3033 English for Occupational Purposes, gave the business environment in which this internship. Macroeconomics and Entrepreneurship pushed the lens even wider, illuminating not just calculation, but the essential calculus of supply, demand, and margins that drive price decisions. The English class was just as important,

emphasizing the ability to communicate ideas effectively. It proved that if people cannot comprehend the mathematical solutions, it may not matter how good they are. These business and communication lessons combined so that it could distill complex data into actionable reports that could be used to inform management decision directly.

3.4 Knowledge Learned that Assisted in Industrial Training

The academic syllabus, described in Section 3.3, can be directly compared to the modes of knowledge gained during the internship. Theoretical knowledge was made less abstract and became rather our basic instrument for engineering real commercial solutions. From manual auditing through system engineering, the development of roles positioned this capacity in service of a system.

The first phase of manual operations and market orientation was rooted in the Economic Principles and communication standards that determined the academic output. By using the theoretical background of "Market Equilibrium," I turned periodic price checks into a systematic competitive analysis of the market, treating the price gaps not just data points but interpreting them as strategic signals of supply and demand instead. The standards for how to communicate in a professional way were taught in the English for Occupational Purposes course. The Microsoft Office Suite (Word, PowerPoint, Excel) was used daily to generate reports based on these standards. It was more important that the information be readable and consumable for the management team, rather than provide raw data.

As the role evolved into data analysis and tool development (Phase 2), most of the solutions were based on the foundation that build in the Statistics and Optimization frameworks studied at university. The Statistics and Advanced Statistics sharpening Mathematical Sensitivity and Business Logic that help in architecting the complex nested logic within the Master Price Comparison Tool. This allowed me able to construct dynamic logic gates that could handle multi-variable scenarios with precision. The most important is the logic that used in the "Shopee Product Strategy Analyzer" to find the optimal balance between ad spend and sales volume. It was not a heuristic guess, but a practical implementation of Constrained Maximization problems derived

from Optimization modules. This guaranteed that marketing spent was being distributed with same budget but return with high-ROAS offerings.

The coursework determined the Computational Rigor and Stochastic Modelling principles that went into building the system architecture in the final phase of comprehensive system automation (Price Intelligence Hub V3). The solid python backend logic that transformed from simpler script was build based on the coding logic learned from Programming Language (C++) class. It provided a sense of how to structure complex classes and functions. In the process of designing the web scraper, the solution to the network instability, the “Retry Logic” was discovered in the iterative convergence principles of Numerical Analysis. The "Human Mimicry" features specifically designed to bypass the anti-bot detection systems were similarly straightforward applications of stochastic random variables. It demonstrating that the theoretical mathematics at the highest level can be bring wise solution to resolve a real-world problem. Moreover, it also proof that it is a need to overcome highly sophisticated modern anti-scraping defence.

3.5 Student's Contribution to the Organization

The company had two different solutions developed to focus on distinct aspects of their data operations. The first of these significant contributions came through the Price Intelligence Hub V3, a hybrid software ecosystem that acts as a continuous monitor on the competitor market. This is a human behaviour mimicking system in the backend consisting of a Python Flask backed logic built with a Chrome Extension scraped seller and competitor data automatically in the background. Now, it used to take three continuous days of donkey work just to perform a manual price comparison for 400 mere products. With this setup, 100% coverage of active products is achieved with minimal human error and this entire system now runs on its own. Additionally, as a self-service platform, the sales team is able to go into the system on their own, run comparisons, and not have to depend on manual data requests.

Building on the competitor tracking, individual tools were created to focus on internal company sales performance: the Shopee & Lazada Product Strategy Analyzers. These

Streamlit applications then take the raw data files from the company and filter them for distinct revenue possibilities. These tools utilize a "Scale Winner" logic to identify products with a high ROAS that are ready to be scaled with ads, and "Engagement Remediation" algorithms to detect listings that receive clicks but fail to convert. The marketing team then reviews the lists to choose on which product should be promote more aggressively, or suggest retargeting plan as needed.

3.6 Organization's Contribution to the Student

The organization afforded an opportunity for the type of practical learning and achieving developmental milestones. There was a fair amount of freedom over which tools and methodologies to use for each task. This autonomy allowed for more engagement with the work and encourage solving problems rather than just mechanically going through the fix routine.

In terms of resource availability, the company supported the need for operations staff accordingly. They were given server access to develop automation, and software requests were approved immediately. On the rare occasions that technical challenges arose, it was the senior staff and the supervisor who provided guidance to me in way that issues were resolved quickly and accurately.

The other benefit was working with real business data. In fact, the internship used real-price competitor prices and real sales data. An example is the advertising budget for the 11.11 campaign was partially passed down for me to direct management. It meant real money, real consequences, which is the way that the university projects didn’t teach me.

3.7 Communication Networks Established

During the internship, there was constant communication with various team leads to update each other on progress. Progress was tracked and issues resolved via frequent check-ins with the supervisor. This included breaking down technical details for non-

coding staff. One of the aims has always been to ensure that the logic of the system is transparent to all.

From a software perspective, modules exchanged data through a single database. This eliminated the manual handling of files. An automatically-generated weekly report containing pricing updates and action items is also sent to the sales team.

Documentation for each tool was authored to assist in knowledge transfer. They include how to use it, how to fix problems, and how to perform updates, to enable future users to keep the system running.

CHAPTER 4

PROBLEMS ENCOUNTERED

4.1 Introduction

This part details discuss about the actual technical issues faced in CG Ikhlas. Market data in the real world is not identical to school homework. Data on the web is often dirty and unstructured. That job was a lot more than just writing Python code. That also involved diagnosing system crashes, and ensuring the database wouldn't freeze up because it received too much data, and explaining tech issues to the rest of the team.

That was merely the beginning, the first script to execute. The bigger challenge was running it in production. There are many daily challenges such as ant-bot blocks, large files choking the server, random data errors. These were not academic questions: they were real issues that were stopping the Price Intelligence Hub V3 working and needed resolving immediately.

4.2 Skills to be Improved

In addition to programming basics taught in the university, other advanced technical skills were highlighted as areas for development for delivering industry-quality software. The experience also highlighted a gap between the academic world and the actual field of work where handling large systems and modern web automation was far different than that learn in the classroom especially experienced by intern in CG Ikhlas.

The first was limited exposure to Data Mining and Unstructured Data. This was a huge stumbling block. At university, the assignments always made use of some "clean" dataset with a known structure that had been presented to us by a lecturer. On the other hand, the industry needed to collect raw intelligence from the e-commerce sites which were typically unstructured, messy and changing continuously. I need make sure data is right, clean it up, and extract data from non-standard sources was also not something being taught in school. This came with a significant self-taught learning experience in pursuing web-scraping and data parsing tools.

Second, the lacking sense in Database Management and Scalability was another aspect that I need to improve. At the start, the system stored data in Excel files. But when the dataset expanded to tens of thousands of rows, the read/write operations got really sluggish and buggy. That meant having to move to SQLite which is great for small-scale internal use, but as soon as concurrent scraping starts, it runs into "Database Locked" errors. Management is also looking at plans to migrate this to a more robust multi-user setup on a MySQL company server.

Third, the gap in Workflow Automation Skills was evident. While university courses covered programming logic (often in C++), they strictly focused on mathematical algorithms rather than practical business automation. There was no exposure to automating daily tasks, such as programmatically reading Excel reports, renaming thousands of files, or interacting with web forms. Transitioning from "manual data entry thinking" to "algorithmic automation thinking" required learning Python's automation libraries (Pandas, Selenium) from scratch, which are far more relevant to modern workplace efficiency than legacy languages.

Fourth, the most challenging part that take longer development timeline is to bypass Advanced Anti-Bot Security. E-commerce platforms like Shopee and Lazada employ advanced CAPTCHA and bot detection systems that instantly block standard scrapers. The most of the time on internship was spent exploring various extraction methods, which was from standard Selenium, to undetectable browser drivers, to desperately find a solution that could persistently bypass these blocks without getting the company

account banned. This process of trial and error was time-consuming but necessary to ensure the scraper's viability.

Action taken:

- Self-discovered the Python libraries such as Pandas (data cleaning) and Selenium (web interaction), acting as a link between university theory and the requirements necessary for exploitation of real-world data.
- Migrated the system away from simplistic Excel sheets to an engineered SQLite database to fix the immediate hang problems, then creating a technical proposal for MySQL server deployment that expands upon the current solution.
- Performed extensive experiments with various browser-automation tools in order to write a scrapper that had humanlike behavior (random pause, natural scrolling), to bypass successfully anti bot protection.
- Wrote automated scripts to handle repetitive tasks, like batch-renaming files or auto-formatting excel reports, in order to turn hours of work into instant programs.

Recommendations for Improvement:

- The university should definitely consider moving the introductory programming language for students in Finance and Applied Math from C++ to Python. In industry Python is used for data analysis, automation and AI while C++ would not really be used by a junior analyst in most modern finance positions. This change would dramatically improve the job prospects and actual tasks for grads.
- Data Mining (on top of open-source tools) and Process Automation to be compulsory, not optional. Finance and Banking students need to understand how to extract data from open sources and automatize recurring workflows - be it just via advanced Excel coding, or Python. Existing manual data entry has obviously seen its day in the modern workplace.
- One of these is a knowledge shift in data storage. More often then not students will bang into a wall when data volume is too big and they simply haven’t been

introduced to databases (SQL/NoSQL). A class on Database Management Systems is necessary, so we're not "clueless" on how to store and retrieve large data set efficiently.

4.3 Communication Skills to be Improved (Technical to Non-Technical)

One of the most challenging of internship period involved translating technical term for colleagues without a tech background. The role demanded more than just writing software; it required turning raw data outputs into usable business intelligence. This was a significant step away from university assignments where the code itself is typically the deliverable for a lecturer that has intimate knowledge of how it works.

This gap is most clearly show up during the conversations about Ad Performance. More frequently the marketing department asked questions like where to spend the budget? or "It is worth spending money on ads for this product?". Delivering standard excel files with technical columns such as "Impressions" has shown to be useless as these metrics do not cater to the financial bottom line. Hence, a change in the way of reporting was necessary. For example, during the 11.11 Campaign, the report was no longer delivered as an excel document but was briefed verbally on the "Scale Winners" list, so the team knew exactly why these products in particular were suggested to receive budget increment.

The second one was the Price Competitiveness. Before the automation, the team needed to manually check prices to check for accuracy. Even when the Price Intelligence Hub V3 was launched, the marketing team still hesitate to trust their own eyes over the automated results. The solution that I did is visualize the math. The team had to understand that a red flag was not a bug, it was a consequence of the simple rule: Competitor Price < Our Price. By passing them through this logic, the dashboard able to reflected powerful trustable insight rather than manual findings. Such transparency removed the fear of the Black Box, and giving them the confidence to finally retire their manual spreadsheets.

Action Taken:

- With the added benefit of sitting next to making team, it became a matter of direct, face-to-face explanations. Rather than only sending files, the thinking behind the "Scale Winners" and "Engagement Remediation" was explained verbally as the data was produced. The team then knew why a product had been recommended for an increase in budget without having to read through complex technical report.
- As reading raw data is time-consuming, the dashboards were tailored to the user language updates. Rather than printing a normal price in the price column, we programmed the system to print a tag on the price tag itself which says "Overpriced by RM5". This simple change meant the team could instantly spot red flags without needing a calculator.
- Short sessions were held to guide the marketing staff on how to read the new automated reports. The goal was to make them self-sufficient. Once the team saw how simple the process was, they started checking the "Competitor Price" lists themselves. This stopped the constant requests for help whenever a new campaign launched.

Recommendations for Improvement:

- Executive Summary" Assessment in Technical Courses: Ideally, specialized subjects like Financial Forecasting (MKW4023) should integrate a "Managerial Reporting" component. Instead of grading solely on mathematical accuracy, 15-20% of the project marks could be allocated to an "Executive Summary", where students must explain their forecast's business implications (e.g., "Risk of Ad Overspend") to a non-expert audience. This directly mirrors the industry requirement of answering "How much should we spend?" without relying on jargon.
- Introduction of Cross-Disciplinary Electives: Recognizing that the transition to Data Analytics was a personal initiative, it is suggested that the university offer optional electives or workshops on "Applied Data Science for Finance". Unlike mandatory core subjects, these elective modules would allow students with a specific interest in technology to explore tools like Python and SQL early in

their degree. This accommodates students seeking to shift to the expanding Fintech/Data Analyst market without an undue technical burden on those pursuing a traditional financial (mathematics-based) path.

4.4 Operational Challenges & Solutions

Apart from the individual development modules, the deployment stage exposed some operational limitations that needed quick engineering fix to keep the system intact. The first big headache coming up was the “Lazy Loading” on modern e-commerce sites. This is a technique used by websites to reduce bandwidth. They hide pictures until a visitor scrolls down to them. Scraper which was original was too fast. It attempted to screenshot the photo before the site displayed it. The bot was not waiting for the visual elements to load, and it kept filling the database with empty records. This was solved by rewriting the extraction logic to follow an "Anchor-Based Navigation" pattern; the script would scroll directly to the bottom of the page where all network requests would be fired and only then, initiate any data collection attempt, thus making sure 100% of the assets are available.

The other big challenge was Operational Fatigue during volume days like the 11.11 Mega Campaign. Needless to say, using humans to verify thousands of products over time decreased the quality of the data, as operators would become fatigued and start to miss entire rows or pages of listings. A hard-coded "Strict Navigation" constraint to enforce data integrity mechanically in the browser extension. It effectively prevented human error by physically disabling the operational-control interface ("Next Page" button in user interface) until certain validation circumstances were set (e.g., all items need to be checked).

The risk of IP Preservation from automated defense systems (WAF) was another key variable with respect to network security. Ongoing aggressive data collection puts the office network at a significant risk of being marked as malicious, possibly leading to a permanent office wide IP ban and hindering all company business. To avoid this, a defensive protocol of "Kill Switch" was injected into it. Instead of bypassing security issues like CAPTCHAs which is a major risk for escalation, this system listens to the

DOM for security components. Once it notices, the process kills itself and breaks the link, because it thinks that ultimately the safety of the office network is more important than the short-term gathering of data.

The system also had to be hardened against the Ingestion of Malformed External Data. The exports from the third-party platforms (such as Shopee Provided Excel file) were interfacing with often came in a wide variety of broken formats like corrupted headers, for mixed character encodings that resulted in various standard Python data libraries throwing repeated exceptions on even the tiniest of data rows. First, the file ingestion module needed a small refactor to handle these inputs as raw data streams, not structured spreadsheets. The complete data architecture was made resilient to clean and process "broken" files without collapsing the entire data pipeline, as forward-passed those broken files into our data-pipeline without applying the strict schema validation during the read-phase on the first place.

Recommendations for Improvement:

- University (Practical Data Handling): Courses such as Financial Forecasting (MKW4023) provide great approaches for mathematical modelling, but tend to use "textbook" datasets that are precleaned. This results in a mismatch when new grads enter the real world, where data is rarely that perfect. Where it could do better: there is no such thing as a "clean" dataset in the real world, and so why not incorporate 'Data Cleaning' components into these math classes that require students to write code to clean up missing or corrupted values prior to being allowed to run their predictions? It makes sure students are truly resilient in technical terms and not just in theory.
- Company (Infrastructure Risk Management): From an operational perspective, running high-volume data collection on the main office network is a "Single Point of Failure". To mitigate the risk of IP bans affecting daily staff operations (like email or E-commerce Platform), it is recommended to separate scraping traffic onto a dedicated line or VPN. This comprises the basic risk management principle "to isolate high-risk assets" into this environment.

CHAPTER 5

CONCLUSIONS

CG Ikhlas was so much of a turn in my path where this 24-weeks internship helps me connect the dots between what I only learn in my college and what the industry will need from me as part of the analytic professional field. In this role of Data Analyst, I started to build my responsibilities around actual operations, e.g: manually checking the prices of competitors every single day using spreadsheets. However, the role organically evolved. Before I knew how to analyze data, I learned there was no way to do that if the systems to build up the data were not engineered. It was out of this necessity that I evolved from a conventional analyst to a Full-Stack Developer where I went from "doing the work" to engineering a complete system of tools that can do the work faster.

Development of my own suite of software was the turning point into a technical developer. I began developing a dedicated Seller and Competitor Collectors to collect raw market data. After that, I created the Shopee/Lazada Strategy Product Analyzer which will help to maximize ad spend with the potential SKU. Hence, I created a Price Intelligence Hub V3, a hub that bridged every data stream for live observation. I did everything on the job by myself. To troubleshoot the engineering problems that came up I taught myself Python, SQL and web basics (JS, HTML, CSS) with no classes. A perfect example is the Campaign Extractor. After experiencing human errors during the 11.11 Mega Campaign, I wrote a "Strict Navigation" protocol that locks the UI until all safety checks were passed. While solving for these operational bugs confirmed at least one thing, I can do more than just design a math model. It meant I could create the robust infrastructure required to supply it with clean data.

My Financial Mathematics background and practical software engineering worked in synergy. My college courses provided me with the logical foundation, learning how to measure variance and think in terms of quantity. But it was the new technical abilities that enabled me to bring those ideas to life at scale. This internship demonstrated that a modern analyst has to be a hybrid professional. It proof that IT-Retailer companies need a "Consultative Analyst" that able to translate raw, chaotic data into clear business strategies.

This internship was a confirmation for me that my career path works. I was no longer a theoretical model student. I'm now a proficient practitioner able to work through data from its inception all the way to handling live ad budgets. The valuable part of the tooling ecosystem I developed will remain an asset to CG Ikhlas, and was a great starting point for me in the road of Data Science and Engineering.

REFERENCES

Barney, J. B. (1991). *Firm resources and sustained competitive advantage. Journal of Management*, 17(1), 99–120. https://doi.org/10.1177/014920639101700108

Bose, R. (2008). *Competitive intelligence process and tools for intelligence analysis. Industrial Management & Data Systems*, 108(4), 510–528. https://doi.org/10.1108/02635570810868362

Feher, C., Elovici, Y., Moskovitch, R., Rokach, L., & Schclar, A. (2012). *User identity verification through mouse dynamics. Information Sciences*, 201, 19–34. https://doi.org/10.1016/j.ins.2012.02.066

Ferrara, E., De Meo, P., Fiumara, G., & Baumgartner, R. (2014*). Web data extraction, applications and techniques: A survey. Knowledge-Based Systems*, 70, 301–323. https://doi.org/10.1016/j.knosys.2014.07.007

Getoor, L., & Machanavajjhala, A. (2012). *Entity resolution: Theory, practice & open challenges. Proceedings of the VLDB Endowment*, 5(12), 2018–2019. https://doi.org/10.14778/2367502.2367564

Glez-Peña, D., Lourenço, A., López-Fernández, H., Reboiro-Jato, M., & Fdez-Riverola, F. (2014). *Web scraping: technologies and applications. Briefings in Bioinformatics*, 15(5), 731–744. https://doi.org/10.52458/978-93-91842-08-6-38

Kannan, P.K. and Kopalle, P.K. (2001). *Dynamic Pricing on the Internet: Importance and Implications for Consumer Behavior. International Journal of Electronic Commerce*, 5, 63-83. https://doi.org/10.1080/10864415.2001.11044211

Sharda, R., Delen, D., & Turban, E. (2014). *Business Intelligence and Analytics: Systems for decision support*.

APPENDIX A: BACKEND ADMINISTRATION & CONFIGURATION

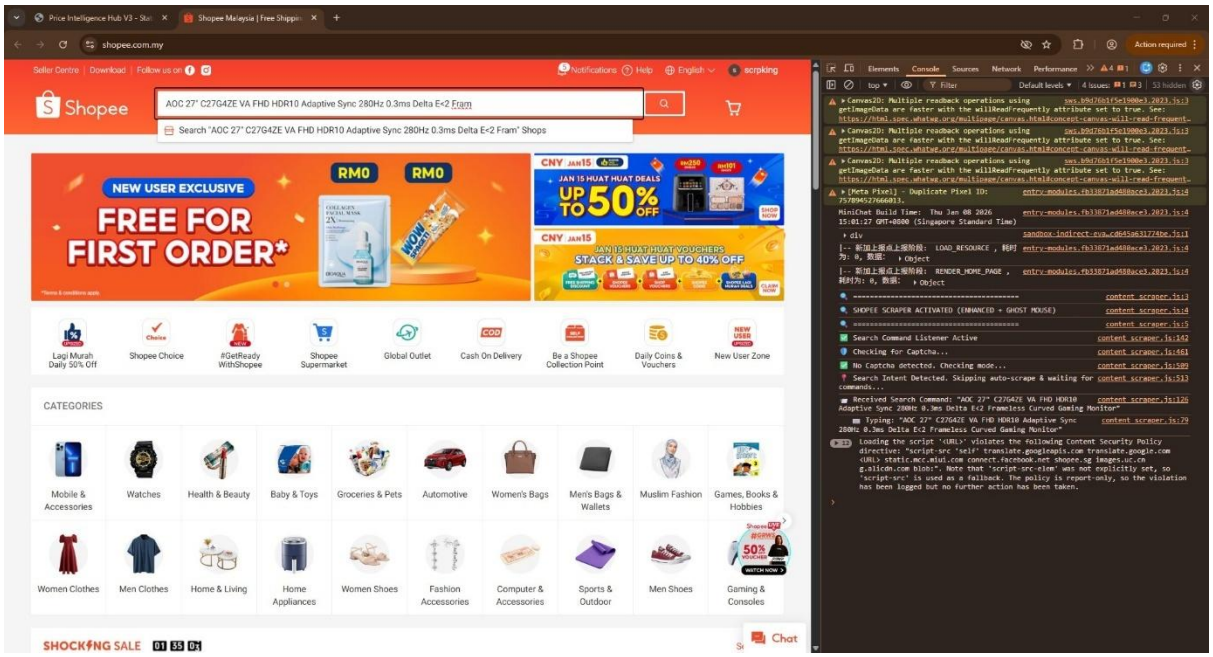


Figure A.1 Search Phase Pattern Injection. The terminal echo output illustrates the "Query Aggregator" engine, which injects its own custom regex patterns (such as ^8GB. *SSD$) to the market search engine and generate the Refined Candidate List.

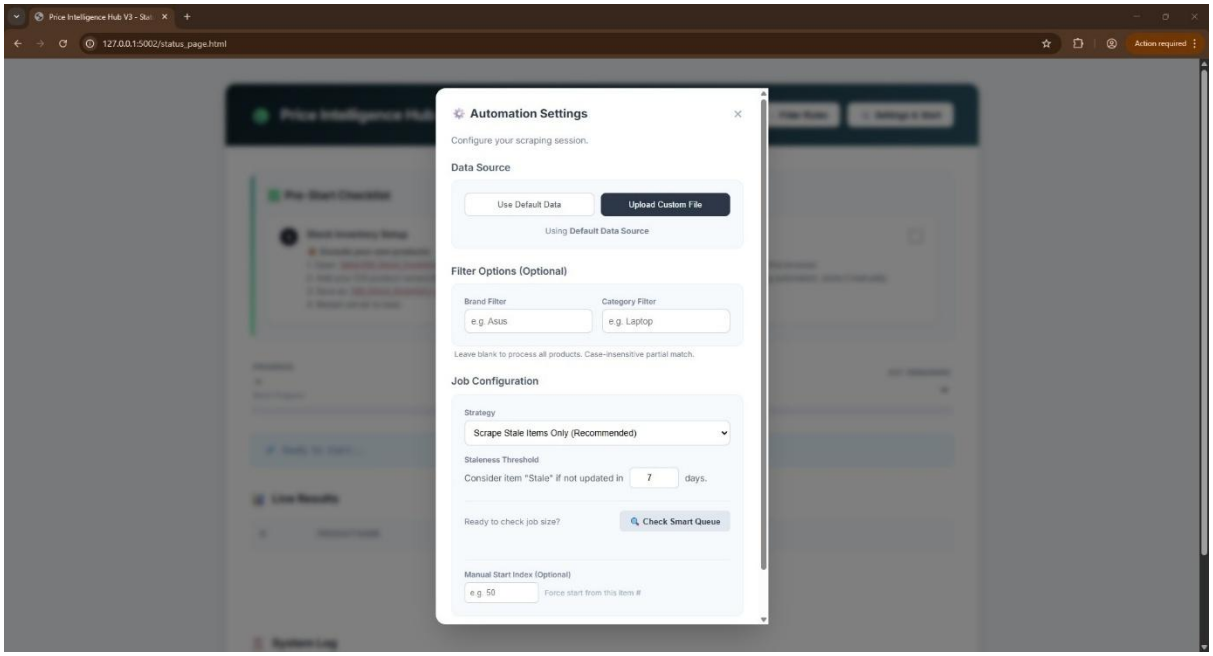


Figure A.2 Operational Batch Configuration. The main interface for configuring the range of daily export. The key features include: 1) Smart Strategy → Whether to process Stale records only (>7 days old and so on), 'Force Scrape' or not; 2) Target Filters: narrow down the list for those high-priority Products on Brands or Categories (for example, "Asus" + "Laptop"); and then 3) Queue Preview which is similar to doubles-checking with a pre-flight calculator showing an estimate of how long this specific job would take at the click of a button.

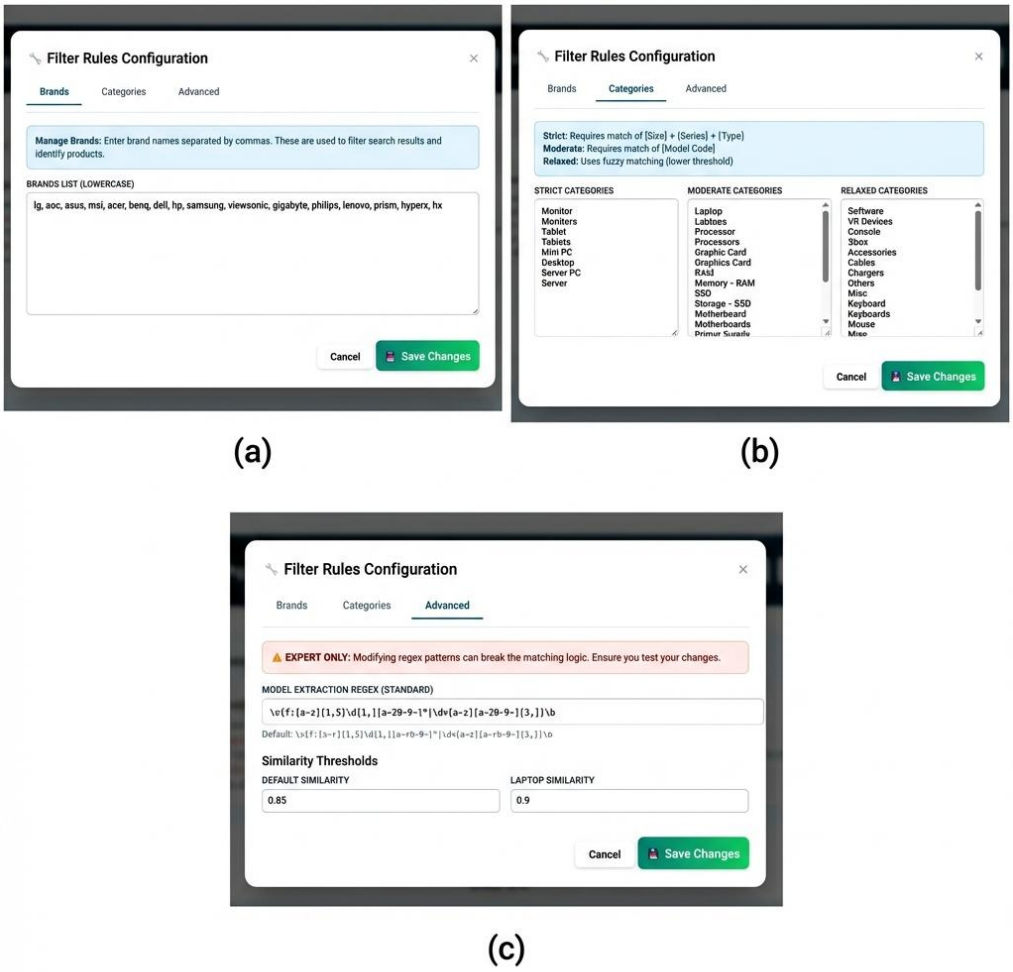
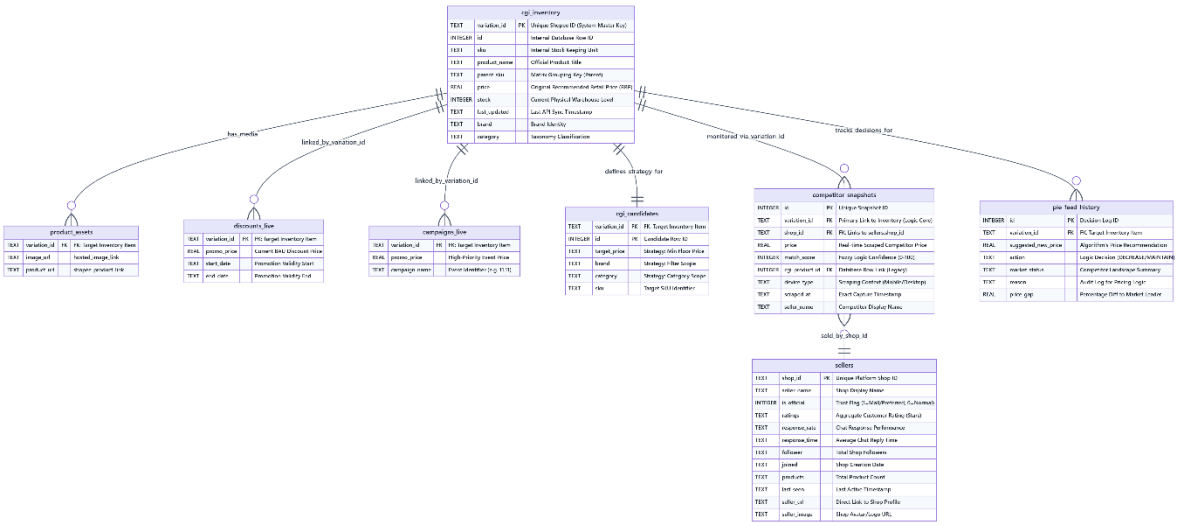


Figure A.3 Multi-Layered Filtering Engine. The proposed system applies a three-level validation pipeline to reduce false positives: (a) Brand Whitelisting: Narrow the search space for known entities, such as ["Asus", "Dell"], to remove brand-generic knock-offs. (b) Taxonomy Levels): Relates categories with Risk-Levels that would usually be more specific (e.g., "Strict" for monitors vs. "Relaxed" for cables). (c) Advanced Fine Tuning: exposes the logic to see the "Glass Box", allows Regex pattern injection and tweak thresholds (default 0.85) for edge cases.

APPENDIX B: DATABASE SCHEMA (SYSTEM ARCHITECTURE)

1) Entity-Relationship Diagram (ERD)

The system is a centralized by "Hub-and-Spoke" database, with the cgi_inventory table being the master immutable record of all pricing, assets and strategy definitions.



2) Data Dictionary (Operational Schema)

Table B-1 Inventory & Strategy Hub

Table	Description	Critical Columns
cgi_inventory	Inventory Master Record. The immutable central registry for all internal products (SKUs), acting as the primary foreign key source for all downstream intelligence operations.	*variation_id (System Master Key)* *sku (Internal Stock Unit)* *parent_sku (Matrix Grouping)* *price (RRP)* *stock (Physical Level)*

cgi_candidates	Search Strategy Scope. Defines the specific "Attack Surface" for daily scraping jobs, mapping internal SKUs to target search keywords (Brand/Category).	*target_price (Min Floor)* *brand (Filter Scope)* *category (Taxonomy)* *variation_id (Target Identity)*
product_assets	Media Abstraction Layer. Decouples heavy binary assets (images) from the core logic tables to optimize query performance.	*variation_id (FK)* *image_url (Hosted Asset)* *product_url (Marketplace Link)*

Table B-2 Competitive Intelligence Layer

Table	**Description**	**Critical Columns**
competitor_snapshots	Market Intelligence Log. A persistent, append-only ledger of all competitor pricing observations captured by scraping agents, serving as the raw evidence for analysis.	*variation_id (Primary Logic Link)* *shop_id (Seller FK)* *price (Captured Value)* *match_score (Confidence 0-100)* *cgi_product_id (Legacy DB Link)*
sellers	Competitor Identity Registry. A deduplicated profile database tracking trust signals, operational metrics, and reputation scores for every unique marketplace seller.	*is_official (Trust Flag: Mall/Preferred)* *response_rate (Chat Performance)* *ratings (Quality Score)* *follower (Social Proof)* *products (Catalog Size)*

pie_feed_history	Algorithmic Decision Journal. A comprehensive audit trail recording every automated price recommendation, the logic rule triggered, and the resultant price gap analysis.	*suggested_new_price (Algo Output)* *action (Decision: Decrease/Maintain)* *reason (Logic Trigger)* *market_status (Context)* *price_gap (% vs Leader)*

Table B-3 Promotion Execution Engine

Table	**Description**	**Critical Columns**
discounts_live	BAU Promotion Layer. Tracks standard, self-managed price reductions (Series Discounts) to distinguish "Net Price" from "Retail Price".	*promo_price (Net Discount)* *start_date (Validity Start)* *end_date (Validity End)*
campaigns_live	Strategic Event Override. Manages high-priority, temporary price overrides for major shopping festivals (e.g., 11.11), taking precedence over standard discounts.	*promo_price (Event Override)* *campaign_name (Event ID)* *variation_id (Target SKU)*

APPENDIX C: CORE ALGORITHMIC LOGIC

i) Human Mimicry Delay

The operational constraint code that prioritizes safety over speed.

```python
def active_wait_mimicry():
    """
    Simulates a human reading the page before navigating.
    Mandatory delay to prevent 'Fast-Scroll' flags.
    """
    # Base reading time (90s) + Variance (30s)
    delay = 90 + random.uniform(0, 30)

    # "Micro-interactions" during wait
    for _ in range(int(delay / 5)):
        scroll_slightly()
        time.sleep(random.uniform(2, 5))
```

Figure C.1 Implementation of the Stochastic Human Mimicry Protocol, utilizing randomized delays and micro-interactions δ_t to evade behavioral bot detection.

ii) Heuristic Matching Engine

The formula used to filter False Positives (Section 2.4.3).

```python
def calculate_match_score(title_a, title_b):
    # 70% Weight: Token Overlap (Keywords match)
    token_score = jaccard_similarity(title_a, title_b) * 0.7

    # 30% Weight: Sequence Order (structure matches)
    sequence_score = difflib.SequenceMatcher(title_a, title_b).ratio() * 0.3

    return token_score + sequence_score
```

Figure C.2 The Heuristic Similarity Algorithm, demonstrating the weighted fusion of Token Overlap (Jaccard) and Sequence Matching logic for candidate scoring.

iii) Specific Use Case Logic: Spec Validation

The "Spec Guard" algorithm that prevents false positives (e.g. 8GB vs 16GB RAM) even if titles match.

```python
def validate_laptop_specs(cgi_prod, seller_name):
    """
    Validate RAM and Storage specs to prevent hardware mismatches.
    """
    cgi_name = str(cgi_prod.get('Product Name', '')).lower()
    seller_name = str(seller_name).lower()

    # 1. RAM Logic (Distinguish 8GB vs 16GB)
    def extract_ram(text):
        # Regex for 1-2 digit numbers followed by 'gb'
        matches = re.findall(r'\b(\d{1,2})\s*gb\b', text)
        return {x for x in matches if x in ['4','8','16','32','64']}

    cgi_ram = extract_ram(cgi_name)
    seller_ram = extract_ram(seller_name)

    # Mismatch if both have RAM info but no overlap
    if cgi_ram and seller_ram and not (cgi_ram & seller_ram):
        return False  # REJECT: Hardware Mismatch

    # 2. Storage Logic (256GB vs 512GB)
    # ... (Similar logic for Storage)

    return True
```

Figure C.3 The Specification "Guard" Logic, acting as a deterministic filter to reject hardware mismatches (e.g., RAM capacity) regardless of title similarity.

APPENDIX D: RAW DATA SAMPLE (INTELLIGENCE OUTPUT)

A raw JSON snapshot generated by the "Robot" during a scraping session, proving the depth of data capture.

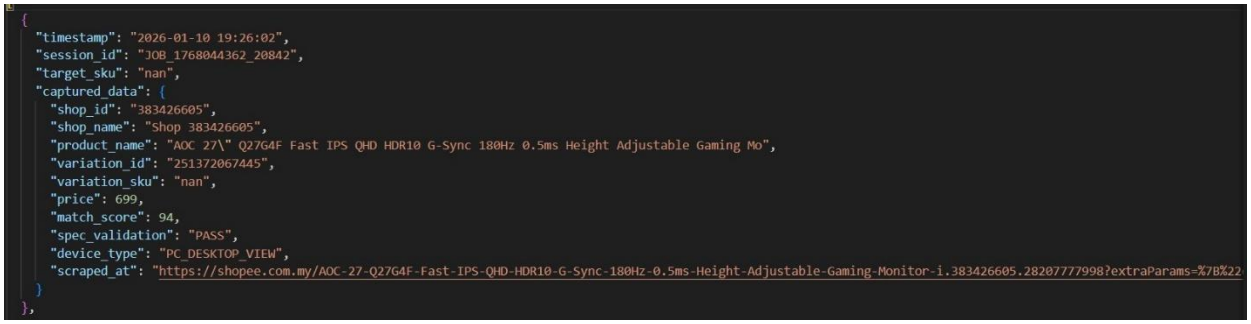


Figure D.1 Live Intelligence Snapshot (JSON) captured from the Production Worker module, showing successful competitor price extraction.

APPENDIX E: LOG BOOK ACTIVITIES

MTM49712

UMT

BUKU LOG LATIHAN INDUSTRI

Minggu 1 (4 Ogos 2025 – 9 Ogos 2025)

Catatan aktiviti setiap minggu

1. Menghadiri sesi orientasi dan penerangan struktur organisasi CGI.
2. Menyemak dan memahami proses perbandingan harga Shopee daripada pelatih sebelumnya.
3. Melaksanakan perbandingan harga manual bagi produk CGI dan merekod tiga pesaing termurah dalam helaian OneDrive.
4. Meneroka kaedah automasi (API & Playwright) bagi mempercepat proses semakan harga.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Kefahaman tentang operasi perniagaan dan struktur produk CGI.
2. Kemahiran dalam menganalisis harga pesaing dan menggunakan formula "Marketability".
3. Pendedahan awal terhadap automasi data menggunakan Shopee API dan skrip Playwright.

Maklumat lain/ Masalah/ Cadangan

1. Proses manual agak memakan masa dan mudah terdedah kepada ralat manusia.
2. API rasmi Shopee memerlukan kelulusan akaun syarikat sebelum boleh digunakan.
3. Cadangan: teruskan pembangunan sistem automasi selepas kelulusan diperoleh.

UNTUK DIISI OLEH PENYELIA INDUSTRI

Catatan

calon menunjukkan ~~minat~~ inisiatif & profesionalism yang tinggi.
laporan disusun dengan kemas & teratur.

Nama, Tandatangan dan Cop Penyelia Industri	
Tarikh:	24/10/2025

CGI IKHLAS SDN BHD

MTM49712

UMT

BUKU LOG LATIHAN INDUSTRI

Minggu 2 (11 Ogos 2025 – 16 Ogos 2025)

Catatan aktiviti setiap minggu

1. Meneruskan kerja semakan harga mingguan untuk produk CGI di Shopee dan mengemas kini helaian perbandingan harga seperti biasa.
2. Mula membangunkan papan pemuka KPI Shopee CGI menggunakan Power BI bagi memantau prestasi jualan.
3. Belajar cara menggunakan data sebenar daripada platform penjual CGI untuk mengira metrik e-dagang seperti jualan bersih, kadar penukaran, dan pembeli berulang.
4. Membangunkan papan pemuka interaktif dengan formula DAX serta perbandingan prestasi (YoY dan MoM).

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Meningkatkan kefahaman tentang metrik utama e-dagang melalui data sebenar daripada platform CGI Seller.
2. Mempelajari cara membina papan pemuka interaktif dan dinamik menggunakan Power BI.
3. Memahami kaedah perbandingan prestasi masa (tahun ke tahun, bulan ke bulan) dalam analisis data jualan.

Maklumat lain/ Masalah/ Cadangan

1. Perlu memastikan formula DAX memberikan hasil yang tepat untuk perbandingan mingguan dan bulanan.
2. Terdapat sedikit kekangan masa antara tugas perbandingan harga dan pembangunan dashboard.
3. Cadangan: Teruskan menambah baik papan pemuka dengan maklum balas daripada penyelia dan pasukan pemasaran.

UNTUK DIISI OLEH PENYELIA INDUSTRI

Catatan

Analisis sangat ~~teliti~~ terperinci & rumalan adalah logik.

Nama, Tandatangan dan Cop Penyelia Industri Tarikh:	24/10/2025 CG IKHLAS SDN BHD

UMT

BUKU LOG LATIHAN INDUSTRI

Minggu 3 (18 Ogos 2025 – 23 Ogos 2025)

Catatan aktiviti setiap minggu

1. Meneruskan kerja semakan harga mingguan bagi produk CGI di Shopee seperti biasa.
2. Menggabungkan data trafik dan prestasi produk ke dalam model Power BI untuk analisis 24 bulan.
3. Menyusun semula halaman papan pemuka Shopee (Traffic → Sales → Product) agar lebih mudah difahami oleh pasukan pemasaran.
4. Membangunkan papan pemuka prestasi iklan Shopee (Shopee Ads Performance) termasuk KPI seperti GMV, ROAS, dan Conversion Rate.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Belajar cara menggabung dan membersihkan data besar (24 bulan) menggunakan Python dan Power BI.
2. Meningkatkan kemahiran dalam reka bentuk papan pemuka e-dagang dengan pelbagai metrik seperti trafik, penukaran, dan pulangan iklan.
3. Memahami hubungan antara trafik, perbelanjaan iklan, dan hasil jualan melalui analisis data sebenar syarikat CGI.

Maklumat lain/ Masalah/ Cadangan

1. Saiz data yang besar menjadikan model Power BI agak perlahan, perlu rancang agregasi data.
2. Papan pemuka Google Sheets mempunyai batasan penapis (slicer) untuk carta; perlu kaji alternatif seperti Looker Studio.
3. Cadangan: Kekalkan tema warna dan susun atur konsisten bagi semua papan pemuka untuk memudahkan bacaan pengguna.

UNTUK DIISI OLEH PENYELIA INDUSTRI

Catatan

Penggunaan sumber data sangat baik.

Nama, Tandatangan dan Cop Penyelia Industri Tarikh:	 24/10/2025

UMT

BUKU LOG LATIHAN INDUSTRI

Minggu 4 (25 Ogos 2025 – 30 Ogos 2025)

Catatan aktiviti setiap minggu

1. Meneruskan kerja semakan harga mingguan bagi produk CGI di Shopee seperti biasa.
2. Dimaklumkan bahawa permohonan Shopee Open Platform API tidak diluluskan, maka tiada cara rasmi untuk automasi penuh semakan harga.
3. Merancang dan mula membangunkan helaian perbandingan harga versi separuh automatik bagi menggantikan proses manual sepenuhnya.
4. Membantu pasukan pemasaran dalam penggambaran iklan kempen Merdeka.
5. Menjalankan kemas kini stok hujung bulan dengan menandakan SKU aktif sahaja dan mengeluarkan item lama atau tidak aktif.
6. Menggabungkan data trafik dan prestasi produk dalam Power BI serta membangunkan halaman *Traffic Overview* dengan penapis peranti (ALL/APP/PC) dan logik "Current vs Previous Period".
7. Menambah kad KPI interaktif (Views, Visitors, Bounce Rate, Avg Time) dengan sistem warna dan ikon dinamik untuk paparan prestasi.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Memahami batasan automasi tanpa akses API rasmi dan kepentingan pendekatan separuh automatik.
2. Belajar mengendalikan kemas kini stok hujung bulan serta mengenal pasti SKU aktif untuk analisis prestasi.
3. Mendapat pengalaman praktikal dalam penggambaran kempen iklan, termasuk penyediaan set dan susun atur produk.
4. Meningkatkan kemahiran membina papan pemuka trafik dalam Power BI menggunakan logik perbandingan masa dan DAX.

Maklumat lain/ Masalah/ Cadangan

1. Kekangan utama ialah ketiadaan API rasmi untuk automasi semakan harga.
2. Cadangan: teruskan pembangunan versi separuh automatik helaian harga untuk mengurangkan masa kerja manual.
3. Pastikan kemas kini stok hujung bulan dilakukan secara konsisten untuk ketepatan laporan produk aktif.
4. Pengalaman membantu dalam penggambaran iklan memberi pendedahan kepada aspek pemasaran visual syarikat.

UNTUK DIISI OLEH PENYELIA INDUSTRI

Catatan

selain menghasilkan analisis yang semakin tepat, calon menunjukkan semangat ~~bekerjasama~~ kerjasama.

MTM49712

Nama, Tandatangan dan Cop Penyelia Industri Tarikh:	24/10/2025 

UMT

BUKU LOG LATIHAN INDUSTRI

Minggu 5(1 Sep 2025 – 6 Sep 2025)
**Nota: 1 Sep 2025 Cuti Ganti Merdeka

Catatan aktiviti setiap minggu

1. Meneruskan semakan harga mingguan bagi produk CGI di Shopee seperti biasa.
2. Membina dan menaik taraf helaian perbandingan harga versi separuh automatik bagi menggantikan proses manual sepenuhnya.
3. Menambah logik warna automatik (green, orange, blue, gray) untuk mengenal pasti isu harga, stok, dan keutamaan tindakan.
4. Membangunkan sistem formula untuk menjana Suggested Price, Priority, Action, dan Reason secara automatik bagi setiap SKU.
5. Menggabungkan semua jenama ke dalam helaian Latest_Combined yang menapis hanya SKU aktif (Active = TRUE).
6. Membina papan pemuka On-Sale Product Health Dashboard dengan visual utama seperti Priority, Gap %, Stock Health, dan Profit Impact.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Belajar membina sistem perbandingan harga separuh automatik menggunakan formula Excel (IFERROR, conditional formatting, Power Query).
2. Memahami kaedah menilai jurang harga (Gap%), keutamaan tindakan, dan kesan keuntungan apabila harga baharu digunakan.
3. Meningkatkan kemahiran dalam visualisasi data melalui pembinaan carta pivot interaktif untuk analisis prestasi jenama dan stok.

Maklumat lain/ Masalah/ Cadangan

1. Tanpa kelulusan Shopee API, automasi penuh tidak dapat dijalankan — helaian separuh automatik menjadi penyelesaian sementara terbaik.
2. Helaian baharu kini membolehkan pasukan mengenal pasti isu harga dengan lebih cepat dan menjana cadangan harga secara automatik.
3. Cadangan: Teruskan penyemakan mingguan dan tambah halaman *drill-through* untuk SKU berisiko atau margin rendah.

UNTUK DIISI OLEH PENYELIA INDUSTRI

Catatan

Pelatihan dalam yang sangat baik.

Nama, Tandatangan dan Cop Penyelia Industri	
Tarikh:	24/10/2025

CG IKHLAS SDN BHD

MTM49712

UMT	BUKU LOG LATIHAN INDUSTRI

Minggu 6 (8 Sep 2025 – 13 Sep 2025)

Catatan aktiviti setiap minggu

1. Meneruskan semakan harga mingguan bagi produk CGI di Shopee.
2. Menambah baik logik helaian perbandingan harga dengan formula baharu yang mengira harga cadangan berdasarkan bilangan pesaing dan jurang harga (Gap %).
3. Menaik taraf sistem Suggested New Price, Action, dan Reason supaya lebih mesra pasukan pemasaran (contoh: "Over by 5 % – restock & adjust").
4. Menetapkan peraturan automatik seperti margin floor, maximum price drop, dan competitor filtering bagi mengelakkan kesilapan harga.
5. Membangunkan helaian Brand Fix Kit, iaitu paparan senarai SKU mengikut tahap keutamaan (Urgent / Medium / Low) untuk setiap jenama.
6. Melakukan pengujian ke atas pelbagai senario (tiada pesaing, harga di bawah kos, stok habis) bagi memastikan formula berfungsi dalam semua keadaan.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Belajar membina formula dinamik Excel (LET, XLOOKUP, FILTER, IFERROR) untuk menghasilkan harga cadangan automatik.
2. Memahami strategi penetapan harga berasaskan pesaing dan margin serta kesannya terhadap keuntungan syarikat.
3. Meningkatkan kemahiran komunikasi teknikal dengan menulis sebab dan tindakan ("Reason / Action") dalam bahasa mudah difahami oleh pasukan bukan teknikal.

Maklumat lain/ Masalah/ Cadangan

1. Disyorkan untuk menambah penapis outlier harga pesaing dan penanda margin keuntungan bagi memastikan cadangan harga lebih tepat sebelum diintegrasi dengan data prestasi iklan.
2. Formula semasa stabil dan sesuai digunakan sebagai asas kepada sistem automatik apabila Shopee API diluluskan kelak.

UNTUK DIISI OLEH PENYELIA INDUSTRI

Catatan

Terdapat sedikit ~~=~~ ruang untuk menambahbaikan dan segi ~~[illegible]~~ ~~penempatan~~. kecepatan.

Nama, Tandatangan dan Cop Penyelia Industri	[signature] CG IKHLAS SDN BHD 201501033748 (1159066-W)
Tarikh:	24/10/2025

UMT BUKU LOG LATIHAN INDUSTRI

Minggu 7 (15 Sep 2025 – 20 Sep 2025)
**Nota: Cuti Am Hari Malaysia (15 Sep & 16 Sep 2025)

Catatan aktiviti setiap minggu

1. Meneruskan kerja semakan harga mingguan bagi produk CGI di Shopee seperti biasa untuk memastikan harga kekal kompetitif.
2. Membangunkan dan menyiapkan papan pemuka prestasi iklan (Ads Performance Dashboards) bagi tiga saluran utama promosi syarikat:
    - Shopee Ads (melalui Meta – FB & IG, link terus ke Shopee Store)
    - Lazada Ads (melalui Meta – FB & IG, link terus ke Lazada Store)
    - IHaveit Website Ads (melalui Meta – FB & IG, link terus ke laman web iHaveit)
3. Menjana laporan KPI utama seperti Spend, Impressions, Reach, CTR, CPC, ROAS, dan Revenue bagi setiap kempen iklan.
4. Menghubungkan data trafik iklan dengan hasil jualan sebenar mengikut siri produk (Expertbook, ThinkPad, LOQ) untuk menilai keberkesanan kempen promosi.
5. Menyusun perbandingan bulan Ogos (sebelum iklan) dan bulan September (selepas iklan) untuk mengenal pasti peningkatan jualan dan prestasi siri produk.
6. Merancang permulaan pembangunan PIE (Pricing Intelligence Engine) bagi menggabungkan data harga, jualan, dan prestasi iklan dalam satu sistem bersepadu.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Belajar menggabungkan data pelbagai platform iklan (Meta Ads dengan pautan ke Shopee, Lazada, dan iHaveit website) ke dalam satu papan pemuka prestasi.
2. Meningkatkan pemahaman tentang metrik utama pemasaran digital seperti CPC, CTR, CPM, ROAS, dan bagaimana prestasi iklan memberi kesan kepada jualan sebenar.
3. Menguasai teknik visualisasi data untuk membandingkan prestasi kempen antara platform serta menentukan saluran paling berkesan.

Maklumat lain/ Masalah/ Cadangan

1. Nilai jualan sebenar daripada kempen Meta yang menghantar trafik ke Shopee dan Lazada tidak dapat dipaut secara langsung, menjadikan metrik ROAS sukar ditentukan secara tepat.
2. Penyelesaian: gunakan jumlah perbelanjaan iklan (Ad Spend) bagi setiap platform dan bandingkan dengan jumlah jualan keseluruhan mengikut siri produk untuk menganggar pulangan (ROAS) dengan lebih realistik.
3. Cadangan: kekalkan semakan harga mingguan dan teruskan pemantauan prestasi iklan secara berkala bagi mengenal pasti siri atau platform dengan pulangan tertinggi.

UNTUK DIISI OLEH PENYELIA INDUSTRI

Catatan

Pecahan data & cerapan yang ~~[illegible]~~ sangat baik.

MTM49712

Nama, Tandatangan dan Cop Penyelia Industri Tarikh:	 24/10/2025 

UMT

BUKU LOG LATIHAN INDUSTRI

Minggu 8 (22 Sep 2025 – 27 Sep 2025)

Catatan aktiviti setiap minggu

1. Meneruskan kerja semakan harga mingguan bagi produk CGI di Shopee seperti biasa untuk memastikan harga kekal kompetitif.
2. Memulakan dan menstabilkan pembangunan PIE (Pricing Intelligence Engine) sebagai sistem analitik berasaskan data untuk cadangan harga automatik.
3. Menghubungkan PIE_FEED, iaitu suapan data yang diambil daripada helaian perbandingan harga (Price Comparison Sheet) — merangkumi hanya produk aktif yang sedang dipantau.
4. Menyepadukan PIE_FEED ke dalam PIE Engine workbook untuk menjadikan ia asas kepada sistem cadangan harga berasaskan data (data-driven suggested price) yang akan dibina pada fasa seterusnya.
5. Menyelesaikan isu teknikal Excel Online seperti #REF! dan #SPILL! dengan membetulkan kawasan tumpahan (spill range) dan menormalkan sambungan data.
6. Menetapkan semula formula utama seperti Floor, Valid_S1–S3, Rank1/2, TargetPrice, Action, Suggested, dan Margin %, serta menambah kawalan ralat (IFERROR, NUMBERVALUE, N) bagi memastikan enjin berfungsi stabil.
7. Menambah lapisan Reason & Action automatik, yang menerangkan secara jelas keputusan enjin untuk setiap SKU (contoh: *"Raised to Floor"*, *"Capped by List"*, *"Hold"*).

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Belajar menghubungkan helaian perbandingan harga sebenar dengan PIE Engine untuk membina sistem analitik bersepadu.
2. Memahami teknik penukaran data teks-ke-nombor (NUMBERVALUE) dan logik pengiraan dinamik (LET, IFERROR) dalam Excel Online.
3. Meningkatkan kefahaman tentang cara mengautomasi pengiraan harga termasuk kawalan margin minimum, pembundaran, dan logik undercut.

Maklumat lain/ Masalah/ Cadangan

1. Excel Online mempunyai batasan pada sambungan external dynamic arrays, yang boleh menyebabkan ralat tumpahan (#SPILL!) dan #CALC! apabila fail dihubungkan terus dari OneDrive.
2. Penyelesaian: pindahkan PIE_FEED ke dalam buku kerja enjin yang sama dan guna fungsi NUMBERVALUE() untuk menukar semua ruangan angka, supaya pengiraan lebih stabil.
3. Cadangan: selepas enjin stabil sepenuhnya, tambah modul untuk pengiraan harga efektif (Effective Price) yang mengambil kira diskaun dan voucher bagi memperhalus cadangan harga masa hadapan.

UNTUK DIISI OLEH PENYELIA INDUSTRI

Catatan

sikap proaktif. konsisten. laporan lengkap & mudah difahami.

MTM49712

Nama, Tandatangan dan Cop Penyelia Industri Tarikh:	24/11/2025 

UMT

BUKU LOG LATIHAN INDUSTRI

Minggu 9 (29 Sep 2025 – 4 Oct 2025)

Catatan aktiviti setiap minggu

1. Menjalankan kemas kini stok hujung bulan (Month-End Stock Update) dengan menandakan SKU aktif sahaja dan mengeluarkan item tidak aktif untuk memastikan data kekal bersih sebelum suapan masuk ke PIE Engine.
2. Meneruskan kerja semakan harga mingguan bagi produk CGI di Shopee seperti biasa untuk memastikan margin dan kedudukan harga kompetitif.
3. Menyemak dan menyegarkan PIE_FEED daripada helaian perbandingan harga (Price Comparison Sheet) agar hanya produk aktif digunakan dalam enjin cadangan harga.
4. Menyediakan dan mengemas kini papan pemuka prestasi iklan (Ads Performance Dashboard) bagi platform Shopee, Lazada, dan iHaveit website (melalui Meta Ads – FB & IG).
5. Membina laporan Proof of Execution (POE) yang merangkumi KPI utama (Spend, Impressions, Clicks, CTR, Purchases, Revenue) serta bukti sokongan seperti screenshot, dan laporan pelaksanaan kempen.
6. Menyiapkan Shopee Dashboard akhir (Aug vs Sep + Phase 1 & Phase 2) sebagai templat utama untuk Lazada dan iHaveit, lengkap dengan demografi (Umur & Jantina), penempatan (Placement), dan siri produk (Expertbook, LOQ, ThinkPad).

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Belajar mengurus data kempen iklan pelbagai platform (Meta Ads → Shopee, Lazada, iHaveit) dan menukarnya menjadi papan pemuka visual yang mudah difahami.
2. Meningkatkan kefahaman tentang penyediaan laporan POE (Proof of Execution) termasuk struktur folder, jadual penjejakan, dan keperluan bukti pelaksanaan iklan.
3. Memahami pentingnya penyelarasan stok dan data aktif sebelum digunakan dalam sistem analitik (PIE Engine dan laporan harga).

Maklumat lain/ Masalah/ Cadangan

1. Terdapat sedikit perbezaan format data antara platform (Shopee, Lazada, Meta), menyebabkan perlu penyesuaian tambahan semasa penyatuan KPI.
2. Penyelesaian: gunakan templat Shopee sebagai struktur asas seragam untuk semua papan pemuka supaya carta, KPI, dan analisis ROAS lebih mudah dibandingkan.
3. Cadangan: teruskan penyelenggaraan stok aktif dan harga mingguan secara konsisten supaya suapan data PIE dan laporan POE kekal tepat serta selaras dengan situasi semasa pasaran.

UNTUK DIISI OLEH PENYELIA INDUSTRI

Catatan

Kerjanya semakin baik ~~[illegible]~~ secara konsisten.

Nama, Tandatangan dan Cop Penyelia Industri	
Tarikh:	04/10/2025

BUKU LOG LATIHAN INDUSTRI

Minggu 10 (6 Oct 2025 – 11 Oct 2025)

Catatan aktiviti setiap minggu

1. Meneruskan semakan harga mingguan bagi produk CGI di Shopee dan kemas kini data aktif untuk suapan PIE.
2. Menyiapkan papan pemuka iHaveit Ads Performance supaya selaras dengan Shopee dan Lazada (struktur KPI, fasa iklan, penempatan, demografi).
3. Menambah analisis Perbelanjaan Iklan mengikut Penempatan dan Platform (FB vs IG) untuk semua papan pemuka.
4. Memulakan penyediaan Proof of Execution (POE) — bina tracker, susun folder, dan kumpul bukti iklan (screenshot, laporan KPI).

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Belajar membina struktur POE lengkap dengan bukti pelaksanaan dan laporan KPI.
2. Memahami analisis prestasi iklan rentas platform (Meta → Shopee, Lazada, iHaveit).
3. Meningkatkan kemahiran penyelarasan data dan visualisasi KPI dalam Power BI dan Excel.

Maklumat lain/ Masalah/ Cadangan

1. iHaveit tiada data jualan dalam platform; POE hanya gunakan metrik iklan (Spend, CTR, CPC) sebagai bukti prestasi.
2. Cadangan: teruskan semakan harga mingguan dan kemas kini PIE_FEED agar data harga dan prestasi iklan kekal selaras.

UNTUK DIISI OLEH PENYELIA INDUSTRI

Catatan

Analisis tajam dan kerja penjelasan teliti.

Nama, Tandatangan dan Cop Penyelia Industri

CG IKHLAS SDN BHD 201501033746 (1159066-W)

Tarikh: 24/10/2025

UMT

BUKU LOG LATIHAN INDUSTRI

Minggu 11 (13 Oct 2025 – 18 Oct 2025)

Catatan aktiviti setiap minggu

1. Meneruskan semakan harga mingguan dan kemas kini stok aktif bagi produk CGI di Shopee.
2. Mengekstrak data promosi diskaun Shopee untuk setiap jenama dan dimasukkan ke dalam modul PIE_Discount bagi membina sistem harga separa automatik yang peka terhadap tempoh promosi (period-aware pricing).
3. Tujuannya ialah supaya sistem PIE dapat mengenal pasti bila promosi sedang berlangsung tanpa perlu mengemas kini harga Shopee secara manual setiap minggu.
4. Menggunakan data promosi ini juga membantu memahami produk sebenar yang dijual oleh CGI merentasi pelbagai jenama dan kategori.
5. Membangunkan dashboard promosi yang memaparkan:
    - Bilangan produk dan kategori dalam promosi,
    - Tempoh promosi (Oktober 2025 – Mac 2026),
    - Jenama dan kategori yang sedang aktif, bagi membantu analisis harga di helaian perbandingan hadapan (Price Comparison Sheet).

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Belajar cara mengautomasikan pengesanan promosi menggunakan data tempoh diskaun (Start–End) untuk meningkatkan ketepatan harga dalam PIE.
2. Memahami cara menstruktur dan memvisualkan data promosi mengikut jenama dan kategori bagi pemantauan yang lebih jelas.
3. Menghubungkan maklumat promosi dengan helaian perbandingan harga untuk menyokong keputusan harga yang lebih pantas dan berasaskan data.

Maklumat lain/ Masalah/ Cadangan

1. Format fail promosi berbeza antara jenama menyebabkan perlu penyesuaian sebelum dimuat naik ke PIE.
2. Penyelesaian: gunakan format templat piawai (PIE_Discount) dengan medan yang seragam (SKU, harga asal, harga diskaun, tarikh mula/tamat).
3. Cadangan: perluas automasi supaya PIE dapat menjejak promosi baharu secara berkala, sekaligus mengurangkan keperluan kemas kini manual mingguan.

UNTUK DIISI OLEH PENYELIA INDUSTRI

Catatan

Kerapian dan perbuatan sila

MTM49712

Nama, Tandatangan dan Cop Penyelia Industri Tarikh:	 24/10/2025

MTM49712

UMT BUKU LOG LATIHAN INDUSTRI

Minggu 12 (20 Oct 2025 – 25 Oct 2025)
**Nota: 20 Oct 2025 (Cuti Deepavali)

Catatan aktiviti setiap minggu

1. Meneruskan semakan harga mingguan bagi produk CGI di Shopee dan mengemas kini helaian Price Comparison setiap minggu.
2. Menjadikan sistem promo-aware: gabungkan data kempen/diskaun (Campaign_Local & Discount_Local) supaya cadangan harga automatik mengambil kira harga promo semasa (jika Promo Status = "Ongoing" dan tarikh dalam julat).
3. Menambah pembersihan data (NUMBERVALUE / TRIM / SUBSTITUTE) dan pengecualian untuk memastikan pengiraan promo-aware stabil.
4. Menyertai penggambaran iklan ASUS ExpertBook — idea keseluruhan video (konsep) datang daripada saya; saya bantu urus set, susun produk dan proofing rakaman.
5. Mengeksport ringkasan pelaksanaan (execution snapshot) dan pautkan bukti iklan ke POE tracker.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Praktikal: menulis formula promo-aware yang stabil bekerja dengan jadual tempatan (Campaign_Local / Discount_Local).
2. Teknikal: pembetulan masalah teks→nombor dan penjagaan ralat (#REF/#SPILL) untuk aliran data berulang.
3. Operasi & kreativiti: pengalaman penggambaran iklan (idea-to-shot), penyediaan produk untuk set, dan komunikasi ringkas dengan pasukan kreatif.

Maklumat lain/ Masalah/ Cadangan

1. Masih guna aliran copy/paste untuk Campaign_Local kerana pautan langsung PIE ↔ Price Comparison tidak stabil di Excel Online.
2. Cadangan operasi: standardkan eksport PIE (CSV) dan jalankan paste/refresh sekali sehari; atau nanti guna Power Query untuk automasi yang lebih stabil.
3. Cadangan kreatif: dokumentasikan idea (shot list + skrip ringkas) supaya sesi penggambaran berikutnya lebih efisien dan boleh digunakan semula untuk kempen lain.

UNTUK DIISI OLEH PENYELIA INDUSTRI

Catatan

calon mengampaikan analisis dengan tepat dan disokong dengan baik.

Nama, Tandatangan dan Cop Penyelia Industri	CG IKHLAS SDN BHD 201501033746 (1159066-W)
Tarikh:	24/10/2025

BUKU LOG LATIHAN INDUSTRI

Minggu 13 (27 Oct 2025 – 1 Nov 2025)

Catatan aktiviti setiap minggu

1. Pembangunan Dashboard Prestasi Iklan Shopee: Mengumpul data CPC dari Shopee Seller Centre dan membina dashboard interaktif untuk menganalisis GMV, ROAS, perbelanjaan iklan, dan keberkesanan kata kunci.
2. Finalisasi Logik Harga Promo v2: Memindahkan rujukan data dari pautan luar ke helaian tempatan (*Discount_Local* & *Campaign_Local*) untuk meningkatkan kestabilan formula dan mengatasi ralat #REF!.
3. Pengemaskinian "Not in Promo List": Menapis helaian perbandingan harga untuk hanya memaparkan SKU yang aktif dan tidak terlibat dalam sebarang promosi bagi memudahkan pemantauan harga produk tetap.
4. Lawatan Penyelia Industri & Pembentangan Projek: Menghadiri sesi lawatan penyelia fakulti dan membentangkan projek mini bertajuk *"Pricing Impact Engine (PIE): Upgrading the Company's Price Comparison System"*.
5. Integrasi Sistem Mass Update Shopee: Menemui modul "Mass Update" di Shopee Seller Centre yang membolehkan pengekstrakan data produk lengkap seperti *Variation ID*, SKU, dan stok untuk disepadukan ke dalam sistem PIE.
6. Automasi Front-End PIE: Menambah baik formula carian menggunakan fungsi LET dan XLOOKUP untuk membolehkan sistem menarik nama produk, ID variasi, dan nama variasi secara automatik berdasarkan data eksport Shopee.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Analisis Pemasaran Digital: Memahami cara menilai keberkesanan strategi pembidaan kata kunci melalui metrik ROAS dan GMV untuk pengoptimuman iklan.
2. Pengurusan Data Lanjutan: Menguasai teknik pembersihan data (seperti SUBSTITUTE dan VALUE) serta penggunaan fungsi Excel yang kompleks untuk membina sistem yang lebih tahan lasak.
3. Sistem Penyelarasan Stok: Mempelajari cara menghubungkan data eksport portal penjual secara terus dengan model analitik dalaman untuk memastikan maklumat ketersediaan produk sentiasa terkini.

Maklumat lain/ Masalah/ Cadangan

1. Isu: Sambungan rentas buku kerja (*cross-workbook*) sering menyebabkan ralat #REF! atau data hilang apabila fail dibuka pada peranti berbeza.
2. Tindakan: Menggunakan helaian rujukan tempatan sebagai penyelesaian sementara dan merancang migrasi ke Power Query atau API untuk kestabilan jangka panjang.
3. Cadangan: Menyeragamkan format eksport CSV bagi modul promosi supaya kemasukan data ke dalam sistem boleh dilakukan secara automatik melalui skrip, sekali gus mengurangkan ralat manual.

MTM49712

UNTUK DIISI OLEH PENYELIA INDUSTRI	
Catatan	
Nama, Tandatangan dan Cop Penyelia Industri Tarikh:	

BUKU LOG LATIHAN INDUSTRI

Minggu 14 (3 Nov 2025 – 8 Nov 2025)

Catatan aktiviti setiap minggu

1. Penambahbaikan Antaramuka Sistem PIE: Melaksanakan proses pembersihan data pada helaian Price Comparison dengan menyingkirkan produk yang telah dinyahsenarai (delisted) bagi memastikan integriti data analisis kekal relevan dan tepat.
2. Standardisasi Visual Analitik: Menyelaraskan logik Conditional Formatting merentasi semua helaian jenama. Ini bertujuan menyeragamkan kod warna bagi penandaan automatik stok habis (out of stock) dan tahap kecemasan harga.
3. Pembangunan Dashboard Iklan 11.11: Membangunkan papan pemuka interaktif untuk memantau prestasi iklan secara langsung sempena persediaan kempen 11.11. Dashboard ini memvisualisasikan metrik kritikal seperti perbelanjaan, klik, Gross Merchandise Value (GMV), ROAS, dan CTR.
4. Pengoptimuman Kata Kunci (ROG): Merangka strategi bidaan untuk produk berprestasi tinggi (seperti ROG Azoth) dengan menetapkan bidaan permulaan RM0.15 bagi memaksimumkan pulangan iklan (ROAS).
5. Automasi Pengekstrakan Data: Membangunkan skrip Python untuk menyelesaikan isu teknikal pada fail muat turun platform yang tidak lengkap, membolehkan proses pengekstrakan maklumat kempen berjalan secara automatik tanpa gangguan.
6. Naik Taraf Sistem 'Product Checker': Mengemaskini helaian semakan agar lebih responsif terhadap status promosi semasa. Sistem kini mampu menarik ID variasi dan butiran harga secara automatik hanya melalui carian ID produk.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Manipulasi Logik Spreadsheet Lanjutan: Mendalami penggunaan gabungan fungsi Excel yang kompleks ($$AND$$, $$OR$$, $$COLUMN$$, $$MATCH$$) untuk membina sistem kawalan paparan yang dinamik dan konsisten.
2. Analitik Pemasaran Digital: Memahami korelasi antara data prestasi iklan dengan data inventori produk sebenar untuk menilai keberkesanan kempen secara holistik.
3. Pembangunan Skrip Automasi: Meningkatkan kemahiran pengatucaraan dalam aspek pengecaman corak data (*data pattern recognition*), membolehkan penghasilan skrip yang boleh diguna semula (*reusable code*) untuk menggantikan proses manual.
4. Pengurusan Visualisasi Data: Mempelajari teknik penggunaan penunjuk visual yang efektif untuk mempercepatkan proses membuat keputusan berkaitan perubahan harga dan stok.

Maklumat lain/ Masalah/ Cadangan

1. Isu Klasifikasi Produk: Mengenal pasti kelemahan pada fungsi *Category Helper* sedia ada yang bergantung sepenuhnya pada *Regex* nama produk. Kaedah ini didapati rapuh dan sering menyebabkan item tersalah klasifikasi sebagai "Uncategorized".
2. Cabaran Integrasi Data: Menghadapi kesukaran menyepadukan data prestasi iklan dengan data produk dalaman disebabkan ketidaksamaan format data (*data format mismatch*) antara platform e-dagang yang berbeza.
3. Cadangan Transformasi Teknikal: Mengesyorkan peralihan daripada penggunaan formula *Regex* yang panjang kepada penggunaan jadual *CategoryMap* yang lebih

efisien. Untuk jangka masa panjang, dicadangkan penggunaan model pembelajaran mesin (TF-IDF + LightGBM) untuk pengelasan produk yang lebih berskala dan tepat. 4. Cadangan Operasi: Mencadangkan penambahan fungsi penyegaran data automatik (*auto-refresh*) pada dashboard analitik bagi memastikan data sentiasa terkini (*real-time*) semasa kempen berlangsung.
UNTUK DIISI OLEH PENYELIA INDUSTRI
Catatan

Nama, Tandatangan dan Cop Penyelia Industri Tarikh:	 12/11/2025

BUKU LOG LATIHAN INDUSTRI

Minggu 15 (10 Nov 2025 – 15 Nov 2025)

Catatan aktiviti setiap minggu

1. Pengurusan Iklan Shopee: Melaksanakan pelarasan harga bidaan (*bid price*) secara berkala untuk memastikan produk berada di kedudukan optimum dalam hasil carian, sambil mempelajari mekanisma sistem ganjaran kredit iklan.
2. Pemantauan Kempen 11.11: Menjalankan pemantauan masa nyata (*real-time monitoring*) bagi prestasi iklan dan perbelanjaan akaun CGI serta ROG di Shopee dan Lazada dengan bajet RM2,000 untuk memaksimumkan pulangan pelaburan (ROI).
3. Strategi 'Retargeting' Berdata: Mengaplikasikan strategi penargetan semula berdasarkan metrik kritikal seperti *Cart Drop-Off Index* (CDI) dan *Interest-to-Order Gap Ratio* (IOGR) untuk mengenal pasti produk yang mempunyai potensi jualan tinggi.
4. Penjenamaan Semula Asus ExpertBook: Menyelaraskan strategi pemasaran siri ExpertBook dengan memfokuskan bajet kepada model berprestasi tinggi (siri B1) bagi membina kepercayaan pelanggan dan mengelakkan *rating dilution*.
5. Pembangunan Alat Analitik: Membangunkan aplikasi dalaman "Shopee Product Strategy Analyzer" menggunakan Streamlit. Aplikasi ini membantu pasukan membuat keputusan automatik (sama ada simpan, sasar semula, atau buang produk) berdasarkan data.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Analisis Metrik E-dagang: Mempelajari formula analitik teknikal untuk mengukur potensi penukaran (*conversion*) produk dengan lebih tepat:
   - $CDI = \frac{Units\ Added\ to\ Cart - Units\ Placed\ Order}{Units\ Added\ to\ Cart}$
   - $IOGR = \frac{Product\ Visitors\ (Add\ to\ Cart) - Buyers\ (Placed\ Order)}{Product\ Visitors\ (Add\ to\ Cart)}$
   - $ERS = (1 - Product\ Bounce\ Rate) \times Product\ Page\ Views$
2. Strategi Bidaan (Bidding): Memahami perbezaan taktikal antara bidaan manual dan automatik serta kesesuaian penggunaannya mengikut objektif kempen (trafik lwn. jualan).
3. Integrasi & Pembersihan Data: Meningkatkan kemahiran pengatucaraan dengan membina fungsi "Data Health Check" dalam aplikasi analitik bagi memastikan tiada ralat pada data sebelum visualisasi dijana.

Maklumat lain/ Masalah/ Cadangan

1. Refleksi Prestasi 11.11: Hasil analisis menunjukkan jualan kempen 11.11 berjaya mencapai sasaran yang ditetapkan (melebihi pencapaian tahun lalu), membuktikan keberkesanan strategi agresif yang digunakan.
2. Pemerhatian Tingkah Laku Pembeli: Mengenal pasti corak trafik memuncak pada waktu makan tengah hari, makan malam, dan tengah malam, yang selari dengan jadual pengeluaran baucar platform.
3. Penyelesaian Isu Teknikal: Mengesan ralat pada pengiraan ERS apabila paparan halaman adalah sifar. Saya mencadangkan penambahan amaran (*alert*) dalam sistem untuk mengelakkan data yang mengelirukan.

4. Cadangan Operasi: Mencadangkan penggunaan fungsi "Bulk Edit" (muat turun pukal) di Lazada untuk mengubah suai harga dan kata kunci secara serentak bagi menjimatkan masa pengurusan operasi.	
UNTUK DIISI OLEH PENYELIA INDUSTRI	
Catatan	
Nama, Tandatangan dan Cop Penyelia Industri Tarikh:	Elise Chong 13/11/2026 CG IKHLAS SDN BHD

UMT

BUKU LOG LATIHAN INDUSTRI

Minggu 16 (17 Nov 2025 – 22 Nov 2025)

Catatan aktiviti setiap minggu

1. Evolusi Sistem 'Shopee Product Strategy Analyzer' (v2): Melaksanakan rombakan komprehensif pada kerangka kerja penganalisis produk. Versi baharu ini merangkumi analisis corong jualan lengkap (*full-funnel analysis*), bermula dari *Engagement Remediation*, *Cart Friction Analysis*, *Interest-to-Order Gap* (IOGR), hingga ke fasa *Checkout Drop-off*.
2. Algoritma Pengesyoran & 'Watchlist': Membangunkan sistem pemantauan khusus untuk menangani kes pinggiran (*edge cases*) seperti "Cart Friction Watchlist" dan "Scale Watchlist". Algoritma pengesyoran turut ditambah baik menggunakan skor metrik berwajaran bagi penglibatan dan potensi penskalaan produk.
3. Visualisasi Corong Jualan (Plotly): Mengintegrasikan perpustakaan Plotly untuk memvisualisasikan aliran prestasi produk secara grafik. Ini memudahkan pemahaman perjalanan produk daripada skor penglibatan (ERS) sehingga menjadi jualan yang disahkan.
4. Pembangunan Alat Hibrid (Browser Extension): Membangunkan pelanjutan pelayar (Chrome/Brave Extension) untuk mempercepatkan proses perbandingan harga Shopee secara langsung dengan pangkalan data dalaman CGI.
5. Integrasi Bahagian Belakang (Python Flask): Membina pelayan belakang (*backend server*) menggunakan rangka kerja Flask untuk menghubungkan data Excel (PIE_FEED) melalui API. Sistem ini dilengkapi algoritma *fuzzy matching* dan normalisasi teks untuk carian produk yang lebih tepat.
6. Mekanisme Pengesanan Anomali Harga: Mengimplementasikan logik keselamatan untuk menanda produk secara automatik jika harga pasaran dikesan tidak normal (25% di bawah atau 20% di atas harga rujukan CGI), bertujuan mengelakkan ralat padanan model atau aksesori.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Seni Bina 'Browser Extension': Mempelajari struktur teknikal Manifest v3 dan teknik suntikan skrip kandungan (*content script injection*) untuk membolehkan interaksi antara pelayar web dan pelayan tempatan.
2. Pengurusan Data Masa Nyata: Memahami perbezaan kritikal antara pengambilan data dinamik (*dynamic fetching*) dan penyimpanan data dalam memori (*caching*), serta kesannya terhadap ketepatan perbandingan harga SKU.
3. Reka Bentuk UX Berinformasi: Menguasai teknik reka bentuk maklum balas visual menggunakan lencana berkod warna (contoh: "CHEAPEST", "RED ALERT") untuk membolehkan pembuatan keputusan pantas tanpa pembacaan teks yang berat.
4. Optimasi Proses Kerja: Berjaya membuktikan peningkatan efisiensi operasi yang ketara, di mana masa analisis bagi setiap produk dikurangkan secara drastik daripada anggaran 5 minit kepada hanya 30 saat.

Maklumat lain/ Masalah/ Cadangan

MTM49712

1. Isu Kecekapan Pemprosesan Data: Semasa menjalankan analisis corong jualan (*funnel analysis*), saya mendapati sistem menjadi perlahan kerana terpaksa memproses ribuan SKU yang sebenarnya sudah lama tiada stok (*out of stock*). 2. Masalah Klasifikasi Jenama: Logik pengesanan jenama semasa (berasaskan *Regex* mudah) didapati kurang tepat, terutamanya bagi sub-jenama seperti "ROG" yang sering tersalah label sebagai "Asus" biasa, atau aksesori yang tersalah kategori. 3. Cadangan Pengurusan "Ghost Products": Untuk minggu hadapan, saya bercadang membangunkan modul pembersihan automatik (*Ghost Product Cleaner*) bagi mengarkibkan data produk yang tidak aktif melebihi 4 minggu. Ini penting untuk mengurangkan beban pengiraan sistem. 4. Cadangan Pemisahan Seni Bina Sistem: Memandangkan kod semakin kompleks, saya mencadangkan untuk memisahkan fungsi pengumpulan data (*Collector*) dan penganalisis (*Analyzer*) kepada dua entiti berbeza pada masa akan datang bagi mengelakkan ralat berantai (*chain errors*).

UNTUK DIISI OLEH PENYELIA INDUSTRI	
Catatan	
Nama, Tandatangan dan Cop Penyelia Industri	Elise Cheng CG IKHLAS SDN BHD
Tarikh:	12/1/2026

BUKU LOG LATIHAN INDUSTRI

Minggu 17 (24 Nov 2025 – 29 Nov 2025)

Catatan aktiviti setiap minggu

1. Migrasi Automasi Data (ETL & VBA): Minggu ini, saya menumpukan usaha pada pemusatan data. Saya telah menggantikan kaedah lama yang menggunakan formula array (*VSTACK/FILTER*) dengan tiga makro VBA khusus untuk menggabungkan data ke dalam helaian *PIE_FEED* dan *Latest_Combined*. Ini menjadikan sistem lebih stabil kerana pemprosesan data kini diasingkan daripada pengiraan formula keuntungan.
2. Integrasi Power Query ("Hot Swap"): Saya telah menukar sumber data promosi daripada salin-tampal manual kepada *Power Query*. Saya menggunakan teknik *"Hot Swap"* di mana jadual baharu dinamakan semula mengikut nama lama, membolehkan formula harga promosi yang kompleks terus berfungsi tanpa perlu ditulis semula.
3. Pembersihan Stok Lama ("Ghost Products"): Saya membangunkan modul penyelenggaraan *Clean_Ghost_Products_Update* untuk mengesan produk yang aktif tetapi tiada stok melebihi 4 minggu. Sistem ini menggunakan logik penjejak mingguan (*lastCheckDate*) untuk memadam data lapuk secara automatik bagi menjimatkan memori sistem.
4. Automasi Pengedaran Produk (Production v1.0): Menyiapkan makro utama *DistributeAndMaintainProducts* yang secara automatik mencipta helaian jenama baharu menggunakan templat "Asus". Saya menambah logik *"Gatekeeper"* yang memastikan helaian baharu hanya dicipta jika produk tersebut benar-benar mempunyai stok, mengelakkan lambakan helaian kosong.
5. Hierarki Pengesan Jenama: Memperkemas ketepatan pengesanan jenama dengan sistem hierarki. Sistem kini mengutamakan sub-jenama perkakasan (seperti ROG, TUF) dahulu, diikuti jenama standard, dan akhirnya kategori perisian atau generik sebagai pilihan terakhir.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Manipulasi Teks VBA (String Manipulation): Saya mempelajari teknik "Nuclear Option" untuk membersihkan simbol *implicit intersection* (@) dalam rentetan formula Excel menggunakan VBA. Ini penting untuk memastikan formula dinamik (seperti LET) berfungsi apabila disalin antara helaian.
2. Strategi Harga Kompetitif: Memahami peralihan kepada strategi harga bebas kos (*cost-independent pricing*). Fokus kini diberikan sepenuhnya kepada persaingan pasaran, di mana harga diselaraskan berdasarkan jurang harga dengan pesaing (contoh: turunkan harga jika jurang >5%).
3. Reka Bentuk UX (User Experience): Meningkatkan pengalaman pengguna dengan menambah bar kemajuan (*Progress Bar*) pada status bar Excel, membolehkan pengguna mengetahui status pemprosesan data masa nyata.

MTM49712

Maklumat lain/ Masalah/ Cadangan
1. Isu Klasifikasi Jenama (False Positives): o Masalah: Produk papan induk Asus tersalah label sebagai "Microsoft" kerana nama produk mengandungi perkataan "Windows 11". Selain itu, perkataan "Headphones" mencetuskan jenama "HP". o Penyelesaian: Saya melaraskan hierarki supaya semakan perkakasan mendahului perisian , dan menggunakan semakan perkataan penuh (*strict whole word check*) untuk jenama pendek seperti HP. 2. Masalah Rujukan Berputar (*Circular Reference*): Templat asal menyebabkan ralat rujukan berputar pada lajur J dan L apabila disalin. Saya menyelesaikan ini dengan mengodkan formula (*hardcoding*) secara terus ke dalam skrip VBA untuk memaksa nilai statik pada sel tersebut. 3. Cadangan Operasi: Disyorkan untuk memantau helaian penjejakan "See me" bagi memastikan tarikh increment berfungsi dengan betul sebelum menjalankan makro pada skala penuh.
UNTUK DIISI OLEH PENYELIA INDUSTRI
Catatan

Nama, Tandatangan dan Cop Penyelia Industri Tarikh:	Elise Chang CG IKHLAS SDN BHD [illegible]

BUKU LOG LATIHAN INDUSTRI

Minggu 18 (1 Dec 2025 – 6 Dec 2025)

Catatan aktiviti setiap minggu

1. Integrasi AI Hibrid (VBA + Python): Membangunkan satu sistem "jambatan" yang menghubungkan Excel VBA dengan model Pembelajaran Mesin (Python). Ini membolehkan sistem mengelaskan ribuan produk stok lama dalam mengklasifikasi jenis produk secara automatik dengan ketepatan 98.63%, menggantikan proses manual yang memakan masa.
2. Migrasi Senibina V3 (Decoupling): Melaksanakan pemisahan struktur sistem yang kritikal di mana "Pengumpul Data" (*Collector*) dan "Penganalisis" (*Analyzer*) kini beroperasi secara berasingan. Ini menjadikan sistem lebih stabil kerana kegagalan semasa mengumpul data tidak lagi akan merosakkan fungsi analisis.
3. Pengurusan Data Sejarah: Membina jadual sejarah baharu dalam pangkalan data yang menyimpan "snapshot" mingguan bagi harga dan promosi. Ini membolehkan syarikat melihat trend perubahan harga lampau, bukan hanya harga terkini sahaja.
4. Penaiktarafan Logik Pengumpul (Stealth Mode): Menggantikan teknik pemasa rawak (*random timer*) dengan logik "Smart Waiting" dan algoritma lengkung Bezier untuk meniru pergerakan tetikus manusia. Ini bertujuan mengelakkan sistem daripada dikesan dan disekat oleh sekuriti Shopee.
5. Pembangunan Papan Pemuka Interaktif: Mencipta utiliti paparan pangkalan data yang membolehkan pengguna menyemak statistik jualan dan menjalankan pembersihan data melalui menu interaktif tanpa perlu menaip kod SQL secara manual.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Seni Bina Perisian (Decoupling): Memahami konsep pengasingan komponen dalam pembangunan perisian untuk mengelakkan ralat berantai (*cascading errors*) dan memudahkan penyelenggaraan kod pada masa hadapan.
2. Simulasi Tingkah Laku Digital: Mempelajari teknik memprogramkan pergerakan tetikus yang natural dan navigasi tab pelayar yang "beringat" (*stateful navigation*) untuk menjadikan bot kelihatan seperti pengguna manusia sebenar.
3. Algoritma Pengecaman Pengepala (Header Hunter): Membangunkan logik yang mampu mengesan lajur data yang betul dalam fail Excel walaupun nama tajuk lajur tersebut tidak konsisten atau berubah-ubah.
4. Logik Sambungan Pintar (Smart Resume): Menguasai teknik pengurusan tugas di mana sistem mampu menyemak pangkalan data sebelum memulakan kerja, membolehkannya melangkau tugasan yang telah siap semalam untuk menjimatkan masa.

Maklumat lain/ Masalah/ Cadangan

MTM49712

1. Isu Kualiti Data (Nama Kedai vs Produk): Sistem kerap tersalah ambil "Nama Kedai" sebagai "Nama Produk" apabila tajuk produk tidak jelas.
    - Penyelesaian: Saya telah mengemaskini logik penapisan untuk mengabaikan teks yang mempunyai corak "Shop [Nombor]" dan mengutamakan pengambilan nama produk daripada URL.
2. Masalah Duplikasi Inventori: Apabila skrip dimuat semula, data inventori menjadi berganda. Saya telah menambah fungsi pembersihan automatik (*DELETE FROM*) sebelum data baharu dimasukkan untuk memastikan rekod sentiasa tepat.
3. Cadangan Akses Jauh: Untuk memudahkan perkongsian hasil kerja dengan pasukan, saya mencadangkan penggunaan teknologi *tunneling* (Ngrok) yang membolehkan pelayan tempatan saya diakses dengan selamat melalui internet.

UNTUK DIISI OLEH PENYELIA INDUSTRI	
Catatan	
Nama, Tandatangan dan Cop Penyelia Industri Tarikh:	 Erika Chong 12/1/2026

BUKU LOG LATIHAN INDUSTRI

Minggu 19 (8 Dec 2025 – 13 Dec 2025)
**Nota: 11 Dec 2025 (Cuti Hari Keputeraan Sultan Selangor)

Catatan aktiviti setiap minggu

1. Automasi Semakan Harga Promosi (Python): Berjaya memindahkan keseluruhan proses semakan harga promosi daripada pergantungan kepada formula Excel yang berat dan kompleks kepada sistem Python sepenuhnya. Sistem baharu ini menggunakan logik padanan tiga peringkat (Variation ID → SKU → Nama) untuk memastikan data promosi dipadankan dengan tepat.
2. Pengekstrakan Data Kempen (Regex): Membangunkan modul pengecaman teks (*Regex Engine*) untuk membaca fail teks kempen Shopee yang tidak berstruktur. Modul ini mampu mengasingkan maklumat penting seperti ID produk dan harga promosi secara automatik daripada timbunan teks mentah.
3. Transformasi Aliran Kerja Strategi (8-Langkah): Menukar proses penjanaan strategi yang sebelum ini bersifat "kotak hitam" (automatik sepenuhnya tanpa pantauan) kepada sistem 8-langkah yang telus. Pengguna kini boleh mengawal dan memantau setiap peringkat, bermula daripada suntikan data, verifikasi jenama, sehinggalah penyimpanan laporan akhir.
4. Automasi Penyelarasan Laporan (Refresh Tracker): Membina skrip pintar yang mengemaskini laporan manual syarikat secara automatik. Sistem ini menangani isu format tarikh Excel dan memastikan saiz jadual berkembang secara dinamik supaya formula sedia ada tidak terganggu.
5. Sistem Bantuan & Status Penyegerakan: Menambah baik antaramuka pengguna dengan tutorial bergambar dan lencana status "Period Sync". Ini memberi maklum balas segera kepada pengguna sama ada tarikh kempen berjaya diekstrak atau masih mempunyai lompong data.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Normalisasi Jenis Data: Mempelajari teknik menangani ketidaksamaan format data antara Excel (yang sering menukar ID kepada nombor perpuluhan/float) dan pangkalan data (teks). Saya membina fungsi penormalan untuk mengelakkan duplikasi data akibat isu ini.
2. Logik Keutamaan Harga (Pricing Algorithm): Memahami cara membina algoritma penentuan harga jualan sebenar dengan menetapkan hierarki keutamaan: Harga Diskaun > Harga Kempen > Harga Stok Asal.
3. Pengurusan Data Sementara (Staging): Memahami konsep penggunaan jadual "Staging" sebagai kawasan transit untuk menyemak kualiti data sebelum ia disahkan dan dimasukkan ke dalam rekod jualan langsung syarikat (Live Tables).

4. Manipulasi Corak Teks (Regex): Meningkatkan kemahiran dalam menggunakan *Regular Expressions* untuk mengekstrak data berstruktur daripada fail teks yang tidak mempunyai format tetap.

Maklumat lain/ Masalah/ Cadangan

1. Isu Notasi Saintifik Excel:
   - Masalah: Excel secara automatik menukar ID produk yang panjang (contoh: 325218...) kepada format saintifik (3.25E+11), menyebabkan data menjadi rosak dan tidak boleh dibaca oleh sistem.
   - Penyelesaian: Saya membaiki kod Python untuk memaksa format teks (*Force-Format Text*) pada lajur ID sebelum data ditulis ke dalam fail Excel.
2. Masalah Data Tidak Lengkap:
   - Masalah: Mengesan bahawa fail promosi mentah dari Shopee tidak mengandungi maklumat kritikal seperti "Jenama" dan "Kategori".
   - Penyelesaian: Saya mengemaskini logik sistem untuk melakukan carian silang (*lookup*) secara automatik ke atas pangkalan data inventori induk untuk melengkapkan maklumat tersebut.
3. Penyelesaian "Ghost Listings": Mengimplementasikan semakan keselamatan untuk memastikan produk yang disenaraikan dalam kempen tetapi tiada stok fizikal (*ghost items*) tidak dipaparkan dalam laporan analisis, mengelakkan kekeliruan strategi.
4. Cadangan Operasi: Disyorkan agar pengguna menggunakan fungsi "Clean Staging" secara berkala untuk membuang draf data lama yang tidak digunakan bagi mengekalkan prestasi pangkalan data yang optimum.

UNTUK DIISI OLEH PENYELIA INDUSTRI

Catatan

Nama, Tandatangan dan Cop Penyelia Industri	
Tarikh:	12/11/2024

BUKU LOG LATIHAN INDUSTRI

Minggu 20 (15 Dec 2025 – 20 Dec 2025)

Catatan aktiviti setiap minggu

1. Tadbir Urus Peraturan Jenama (Brand Governance): Memperkenalkan sistem keutamaan jenama 3-Tahap (Tinggi/Sederhana/Rendah) untuk menyelesaikan konflik penamaan. Ini memastikan jenama utama seperti "Asus" tidak lagi tersalah ganti oleh kata kunci komponen seperti "Intel" atau "GeForce".
2. Pemisahan Alat 'Seller Resolver' (Port 5003): Membangunkan perkhidmatan mikrosistem berasingan (*Standalone Microservice*) khusus untuk mengenali identiti penjual Shopee. Alat ini beroperasi di pelayan (port 5003) yang berbeza daripada pengumpul data utama untuk mengelakkan gangguan trafik.
3. Galeri Imej & Perbandingan Visual: Membina modul galeri interaktif yang membolehkan pengguna membesarkan (*zoom*), menggeser (*pan*), dan membandingkan imej produk pesaing dengan produk inventori sebelah-menyebelah secara langsung di papan pemuka.
4. Kejuruteraan Anti-Bot Lanjutan: Mengimplementasikan logik peniruan manusia (*Human Mimicry*) yang kompleks dalam skrip automasi. Ini termasuk penggunaan formula matematik Bezier Curves untuk pergerakan tetikus yang melengkung secara natural dan corak skrol "Stop-Go" untuk meniru gaya pembacaan manusia.
5. Integrasi Log Masa Nyata (SSE): Membangunkan konsol log langsung menggunakan teknologi *Server-Sent Events* (SSE). Ini membolehkan pengguna memantau aktiviti backend (seperti pelaksanaan makro VBA atau penyegerakan pangkalan data) terus dari pelayar web tanpa perlu membuka terminal.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Simulasi Biometrik Digital: Mempelajari cara memprogramkan pergerakan robotik supaya kelihatan organik dengan memvariasikan kelajuan dan laluan (bukan garis lurus) bagi mengelakkan pengesanan sistem keselamatan e-dagang.
2. Seni Bina Mikrosistem: Memahami konsep menjalankan aplikasi modular pada port pelayan yang berbeza (*Decoupled Architecture*) untuk meningkatkan kestabilan dan memudahkan penyahpepijatan.
3. Manipulasi DOM Lanjutan: Menguasai teknik *Anchor-Based Scoping* untuk mengekstrak elemen web yang sukar (seperti gambar profil penjual) dengan mencari elemen "sauh" (butang Chat) terlebih dahulu sebelum bergerak ke elemen sasaran.
4. Pengurusan Keutamaan Data: Mempelajari teknik menyusun logik pemprosesan data berdasarkan hierarki kepentingan (Jenama Utama vs Komponen) untuk meningkatkan ketepatan klasifikasi produk.

Maklumat lain/ Masalah/ Cadangan	
1. Isu Ketepatan Imej: o Masalah: Pengikis web (*scraper*) sering tersalah ambil gambar produk sebagai gambar profil penjual pada halaman yang menggunakan *lazy-loading*. o Penyelesaian: Saya mengubah logik carian kepada teknik "Anchor-Based", di mana kod mencari butang "Chat Now" dahulu sebagai titik rujukan sebelum mengambil gambar yang berada dalam bekas (*container*) yang sama. 2. Risiko Sekatan Keselamatan (CAPTCHA): o Masalah: Aktiviti automasi berisiko tinggi menyebabkan akaun syarikat disekat jika CAPTCHA muncul. o Penyelesaian: Melaksanakan protokol "Nuclear Option", di mana sistem akan mati serta-merta (*emergency kill switch*) jika sebarang elemen CAPTCHA dikesan, dan memaksa campur tangan manual. 3. Cadangan Dokumentasi: Memandangkan sistem semakin kompleks, saya mencadangkan dan telah memulakan inisiatif memisahkan dokumentasi kepada tiga bahagian khusus: Panduan Pemasaran (untuk pengguna harian), Panduan Teknikal (untuk pembangun), dan Persediaan Sistem.	
UNTUK DIISI OLEH PENYELIA INDUSTRI	
Catatan	
Nama, Tandatangan dan Cop Penyelia Industri Tarikh:	Eliff Chong …G IKHLAS SDN BHD 12/1/2026

UMT

BUKU LOG LATIHAN INDUSTRI

Minggu 21 (22 Dec 2025 – 27 Dec 2025)
**Nota: 25 Dec 2025 (Cuti Krismas)

Catatan aktiviti setiap minggu

1. Pembangunan Alat Pengumpul Data Kempen: Membangunkan satu sistem baharu untuk menarik data produk kempen Shopee secara automatik. Fokus utama adalah untuk menggantikan proses manual yang memakan masa dengan fungsi "auto-sync" yang mampu merekodkan maklumat produk setiap 3 saat.
2. Sistem Semakan Kualiti Data (Approval Workflow): Memperkenalkan langkah keselamatan baharu dalam proses pengumpulan data di mana data yang dikumpul tidak terus disimpan ke dalam pangkalan data. Sebaliknya, data diletakkan dalam "Senarai Semakan" (*Pending Review*) terlebih dahulu untuk disahkan oleh pengguna, bagi memastikan hanya data berkualiti disimpan.
3. Penambahbaikan Antaramuka Pengguna (UI/UX): Mengemas kini papan pemuka (*dashboard*) sistem kepada susun atur 3-lajur yang lebih kemas (Senarai Tugas, Semakan Tertunda, & Sejarah Disahkan). Ini memudahkan pengguna memantau status kerja dengan lebih jelas berbanding rekaan lama.
4. Automasi Pengurusan Pengecualian: Menambah fungsi pintar yang membolehkan sistem mengesan kedai yang telah "disekat" (*banned*) oleh Shopee secara automatik. Sistem kini akan melangkau kedai tersebut tanpa tersekat, membolehkan proses pengumpulan data berjalan lancar tanpa gangguan.
5. Optimasi Eksport Laporan Pukal: Menambah baik fungsi eksport laporan dengan menyertakan pautan terus (*hyperlink*) ke kedai penjual dan nota statistik terperinci pada sel Excel. Ini memudahkan pasukan pengurusan membuat semakan silang tanpa perlu mencari pautan secara manual.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Integriti & Validasi Data: Mempelajari betapa pentingnya mewujudkan lapisan "staging" (tempat simpanan sementara) sebelum data disimpan secara kekal. Ini mengelakkan pangkalan data daripada tercemar dengan maklumat sampah atau duplikasi.
2. Reka Bentuk Berpusatkan Pengguna: Belajar cara mereka bentuk aliran kerja yang lebih intuitif, seperti menambah butang bantuan (*Help Guide*) dan notifikasi status untuk membantu pengguna bukan teknikal memahami cara menggunakan alat ini.
3. Komunikasi Antara Sistem: Memperoleh kemahiran teknikal dalam menghubungkan pelayar web (Chrome Extension) dengan pelayan data tempatan, serta cara mengatasi halangan keselamatan pelayar yang menyekat sambungan tersebut.
4. Pengurusan Data Dinamik: Mempelajari teknik mengambil data daripada laman web yang kompleks (*Single Page Application*) dengan menggunakan kaedah pemantauan

URL pintar bagi memastikan sistem bertindak balas pantas terhadap perubahan halaman.
Maklumat lain/ Masalah/ Cadangan 1. Masalah Keciciran Data Manual: Semasa mengumpul data kempen, pengguna sering terlepas pandang atau melangkau halaman produk, menyebabkan laporan akhir tidak lengkap. o Penyelesaian: Saya telah membina fungsi "Kawalan Navigasi Ketat" yang menghalang pengguna daripada melangkau halaman. Sistem akan memaksa pengguna bermula dari halaman 1 dan bergerak mengikut urutan. 2. Isu Teknikal Laman Web Moden: Mengesan masalah di mana sistem "tergantung" apabila laman web Shopee bertukar halaman tanpa memuat semula sepenuhnya (*Single Page Application*). Saya menyelesaikannya dengan memprogramkan sistem untuk memaksa muat semula (*force reload*) bagi memastikan skrip berjalan lancar. 3. Cadangan Keselamatan: Disyorkan agar pasukan menggunakan profil pelayar yang berasingan dan VPN statik semasa menjalankan aktiviti pengumpulan data bagi mengurangkan risiko akaun syarikat disekat oleh platform.
UNTUK DIISI OLEH PENYELIA INDUSTRI
Catatan

Nama, Tandatangan dan Cop Penyelia Industri Tarikh:	Elist Chong CG IKHLAS SDN BHD 13/1/2026

BUKU LOG LATIHAN INDUSTRI

Minggu 22 (29 Dec 2025 – 3 Jan 2026)
**Nota: 1 Jan 2026 (Cuti Tahun Baru 2026)

Catatan aktiviti setiap minggu

1. Pemulihan Fungsi Analisis Pukal: Membaiki sepenuhnya halaman "Bulk Analysis" yang sebelum ini tidak berfungsi. Pasukan kini boleh memuat turun data sejarah dan memulakan analisis baharu dengan lancar.
2. Penyelesaian Masalah Eksport Data Besar: Mengubah cara sistem menjana laporan Excel. Sebelum ini, eksport data yang besar akan menyebabkan sistem terhenti (*freeze*). Saya telah memindahkan proses ini ke latar belakang (*background processing*), membolehkan pengguna terus menggunakan sistem sementara laporan disiapkan.
3. Integrasi Galeri Imej Visual: Membangunkan fungsi galeri imej interaktif pada halaman semakan. Ciri ini memudahkan staf untuk mengesahkan produk pesaing secara visual (dengan fungsi *zoom* dan *pan*) tanpa perlu membuka pautan luar satu persatu.
4. Automasi Laporan Stok: Menambah baik skrip automasi Excel untuk menangani fail laporan daripada Shopee. Saya memastikan formula pengiraan dan format jadual disalin secara automatik setiap kali data stok dikemaskini, menjimatkan masa penyuntingan manual.
5. Penambahbaikan Penapis Analisis: Mengemas kini fungsi carian pesaing dengan menambah pilihan "Hide High-Risk". Fungsi ini secara automatik menyembunyikan kedai rakan kongsi dalaman daripada senarai, supaya penganalisis hanya melihat pesaing sebenar sahaja.
6. Alat Diagnostik Pangkalan Data: Membina satu alat penyelenggaraan khas yang membolehkan saya mengeksport keseluruhan pangkalan data ke dalam Excel dengan satu klik. Ini sangat membantu apabila saya perlu mencari punca masalah data dengan cepat.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Pengurusan Proses Latar Belakang: Mempelajari cara mengendalikan tugas berat (seperti menjana fail Excel besar) di belakang tabir supaya tidak mengganggu kelancaran antaramuka pengguna.
2. Pengendalian Format Data Luar: Menimba pengalaman dalam menyelesaikan isu keserasian fail Excel yang dijana oleh pihak ketiga (Shopee). Saya belajar teknik untuk memintas ralat validasi supaya fail tersebut boleh dibaca oleh sistem kami.
3. Penyusunan Kod Modular: Mempelajari kepentingan menyusun semula kod program yang panjang kepada bahagian-bahagian kecil (*modules*). Ini menjadikan sistem lebih mudah diselenggara dan kurang risiko ralat pada masa hadapan.

4. Logik Pemadanan Pintar: Membangunkan logik carian yang lebih "bijak" untuk mengesan produk pesaing walaupun kod produk (SKU) mereka tidak ditulis dengan format yang sempurna atau lengkap.

Maklumat lain/ Masalah/ Cadangan

1. Masalah Kestabilan Sistem: Sistem kerap mengalami ralat "Database Locked" apabila dua proses cuba menulis data serentak.
    - Penyelesaian: Saya telah menyusun semula urutan arahan dalam kod untuk memastikan satu proses selesai sepenuhnya sebelum proses seterusnya bermula, menyelesaikan masalah "berebut" sumber.
2. Isu Fail Laporan Shopee: Skrip automasi kami gagal membaca fail Excel yang dimuat turun dari Shopee kerana format dalaman fail tersebut tidak standard.
    - Penyelesaian: Saya telah membuat pengubahsuaian kecil pada kod pembaca Excel untuk mengabaikan pemeriksaan format yang terlalu ketat, membolehkan fail tersebut diproses seperti biasa.
3. Pembersihan Data: Saya mendapati banyak rekod produk yang hilang gambar atau pautan. Saya telah menjalankan satu skrip pembersihan untuk membuang data yang rosak ini supaya paparan di sistem sentiasa kemas dan tepat.
4. Cadangan Penambahbaikan: Saya mencadangkan agar semua fail hasil eksport disimpan dalam folder yang boleh diakses terus melalui pelayar web (*static folder*), bagi memudahkan staf memuat turun fail tanpa masalah akses.

UNTUK DIISI OLEH PENYELIA INDUSTRI

Catatan

Nama, Tandatangan dan Cop Penyelia Industri Tarikh:	Firo charq CG IKHLAS SDN BHD [illegible]

MTM49712

UMT

BUKU LOG LATIHAN INDUSTRI

Minggu 23 (5 Jan 2026 – 10 Jan 2026)
**Nota: Catat CUTI/CUTI SAKIT berserta tarikh sekiranya CUTI/CUTI SAKIT

Catatan aktiviti setiap minggu

1. Meneruskan operasi rutin pengumpulan data harga pesaing dan maklumat penjual (competitor & seller data acquisition) bagi memastikan pangkalan data pasaran sentiasa dikemas kini untuk analisis mingguan.
2. Melaksanakan penyelenggaraan berkala (system maintenance) pada platform "Price Intelligence Hub" (PIH V3), termasuk pembaikan ralat integrasi aset digital dan pengoptimuman algoritma penapisan "CGI Partner".
3. Menyelenggara dan mengemas kini dokumentasi teknikal sistem secara berterusan bagi memastikan setiap perubahan pada logik kod dan seni bina pangkalan data direkodkan dengan tepat.
4. Menyediakan dokumentasi Standard Operating Procedure (SOP) penyerahan tugas (handover) yang komprehensif bagi memudahkan peralihan sistem kepada kakitangan tetap dan menjamin kelangsungan operasi jabatan.
5. Menyiapkan draf laporan teknikal projek mini (mini project report) yang merangkumi analisis impak perniagaan dan metrik kecekapan operasi untuk semakan akhir oleh pengurusan jabatan.
6. Mengemas kini fungsi eksport pelaporan pada sistem untuk menyokong keperluan pengauditan baru serta memastikan integriti fail hasil eksport dalam folder statik sistem.
7. Melakukan analisis "Kesan Multiplier" bagi membuktikan peningkatan kapasiti pemprosesan data automasi syarikat daripada 40 SKU kepada lebih 450 SKU sehari.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Pemantapan kemahiran dalam penyelenggaraan sistem perisian (software maintenance) dan pengurusan kitaran hayat pembangunan sistem (SDLC).
2. Keupayaan dalam menyediakan dokumentasi teknikal korporat (SOP) yang mematuhi standard profesional bagi tujuan penyerahan projek.
3. Kemahiran menganalisis data prestasi sistem untuk menghasilkan laporan impak teknikal yang menyokong keputusan perniagaan Syarikat.

Maklumat lain/ Masalah/ Cadangan

1. Masalah: Keperluan untuk memastikan dokumentasi teknikal selari dengan kemas kini kod sistem yang dinamik bagi mengelakkan kekeliruan semasa proses penyerahan tugas (handover) kelak.
2. Tindakan: Penyelenggaraan dokumentasi dilakukan secara selari dengan pembaikan pepijat (bug fixes) dan pengoptimuman sistem.
3. Cadangan: Disarankan agar sesi taklimat SOP dijalankan kepada kakitangan berkaitan sebelum tempoh latihan industri berakhir bagi memastikan proses penyerahan berjalan lancar.

MTM49712

UNTUK DIISI OLEH PENYELIA INDUSTRI	
Catatan Prestasi yang amat memuaskan.	
Nama, Tandatangan dan Cop Penyelia Industri Tarikh:	Evie Chan CG IKHLAS SDN BHD 13/1/2024

BUKU LOG LATIHAN INDUSTRI

Minggu 24 (12 Jan 2026 – 17 Jan 2026)

Catatan aktiviti setiap minggu

1. Menstabilkan sistem Price Intelligence Hub V3 (PIH V3) untuk fasa pengeluaran (production-ready). Ini termasuk mengoptimumkan pengindeksan pangkalan data SQLite bagi mengelakkan ralat "Database Locked" semasa aktiviti pengikisan data (scraping) berkelajuan tinggi.
2. Menyediakan dokumentasi Prosedur Operasi Standard (SOP) teknikal yang komprehensif untuk proses penyerahan tugas (handover) kepada pelatih seterusnya. Manual ini merangkumi persediaan persekitaran Python (Pandas, Selenium, Flask), konfigurasi VPN, dan logik akaun "Burner".
3. Merumus keseluruhan pengalaman latihan industri selama 24 minggu ke dalam Laporan Akhir Latihan Industri. Fokus diberikan kepada evolusi peranan daripada audit manual kepada kejuruteraan sistem, serta cara menyelesaikan masalah data tidak berstruktur menggunakan formula padanan heuristik.
4. Menyediakan slaid pembentangan akhir yang memaparkan seni bina 5-fasa projek dan impak komersial sistem (peningkatan kapasiti daripada 40 SKU kepada 450+ SKU sehari) dan keseluruhan praktikal industri.
5. Melaksanakan sesi tunjuk cara akhir (walkthrough) dashboard PIH V3 bersama pengurus pemasaran dan melengkapkan proses pelepasan (clearance) latihan industri.

Pengetahuan/ kemahiran yang diperoleh (sepanjang minggu)

1. Kemahiran dalam menyediakan dokumentasi teknikal korporat (SOP) yang mematuhi standard industri bagi menjamin kelangsungan sistem.
2. Keupayaan untuk menukar skrip pembangunan kepada sistem yang stabil dan sedia untuk kegunaan pengeluaran (production-ready logic).
3. Kemahiran komunikasi strategik dalam menterjemah kod teknikal yang kompleks kepada hasil strategi perniagaan yang boleh difahami oleh pihak pengurusan.
4. Pemahaman mendalam mengenai kitaran hayat data yang lengkap, bermula daripada pengumpulan data mentah sehingga kepada analisis impak jualan (seperti pertumbuhan jualan kempen 11.11 daripada RM70k ke RM120k).

Maklumat lain/ Masalah/ Cadangan

1. Masalah: Memastikan kestabilan sistem untuk jangka masa panjang bagi pengguna bukan teknikal selepas tempoh latihan berakhir.
2. Tindakan: Mengintegrasikan protokol "Strict Navigation" dan "Kill Switch" sebagai langkah keselamatan automatik untuk menghalang ralat manusia.
3. Cadangan: Disarankan agar jabatan berkaitan memantau prestasi pelayan pangkalan data secara berkala dan mengikuti manual troubleshoot yang telah disediakan dalam SOP untuk penyelenggaraan masa hadapan.

MTM49712

UNTUK DIISI OLEH PENYELIA INDUSTRI	
Catatan	
Nama, Tandatangan dan Cop Penyelia Industri Tarikh:	Elise Chong CG IKHLAS SDN BHD 13/11/2024

BIODATA OF THE AUTHOR

Name	: Lai Chun Kit
Nationality	: Malaysian
Permanent Address	: No.50, Jalan Bernam 10, Taman Bernam Baru, 35900 Tanjong Malim, Perak
Telephone Number	: 017-5142980
Email	: vaniepersie@gmail.com
Education	: Bachelor of Science in Financial Mathematics with Honours
Award(s)/ Achievements	: RASIO 9.0 Data Science Competition 3rd Place RASIO 10.0 Data Science Competition Finalist UMT Innovation 2024 Silver Medal (NomSim: Financial Unity Game) SRD UMT 2025 Finalist Top 72 Anugerah Dekan, Semester 1 - 6