AI Project

Week 1 – Retail Analysis with Cursor AI
(KPI, RFM, Return and Abnormal analysis)

Ken Lin 08/11/2025

Guideline

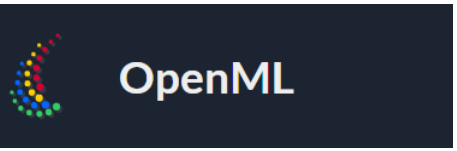
- Project Goal
- Source Data
- Suggested Project Workflow
- Tool Used
- Case Overview
- Case 1 – AI Driven Retail Analytics

Project Goal

- Utilize **AI Cursor** to build automated scripts.
- Automate seamless data pipeline.
- Demonstrate **problem-solving abilities**.
- Showcase **project management skills**.

Source Data

kaggle



Why These Datasets Are Suitable

- Realistic / real-world context (banking and retail transactions).
- Contain **analyzable problems** (customer segmentation, churn, marketing effectiveness, anomaly detection).
- Allow designing a **complete workflow**: data acquisition → cleaning → modeling/analysis → reporting/automation, enabling you to demonstrate end-to-end project execution.
- Can be structured to **update regularly and monitor via dashboards**, highlighting the automation component.

Suggested Project Workflow

Download datasets and quickly review their structure, data volume, and missing values.
Select 1–2 datasets that you are most interested in and believe best showcase your skills (e.g., *Online Retail II*).
Design a project draft for the selected dataset(s), including:

- Project objectives
- Data workflow
- Tool stack (Python / SQL / Power BI / AI Cursor)
- Milestones (project management)

Start building the automated script workflow using **AI Cursor** logic:

- Script initiation → Data acquisition → Cleaning → Feature engineering → Analysis / Modeling → Reporting / Dashboard generation

Build dynamic reports and dashboards and generate actionable data-driven insights to support strategic decision-making and operational improvements.

Tool Used

Case Overview

	Dataset Name	Source Link	Data Volume	Key Features Overview	Project Problem Solved	Status
1	Online Retail II	UCI Repository	1,067,371	Transaction details: invoice, product, quantity, customer	KFM Analysis, KPI Analysis, Anomaly Detection	**Ongoing**
2	Bank Marketing Dataset	UCI Repository	41,188	Demographics + marketing campaign info	Customer targeting & marketing campaign	Pending
3	Telco Customer Churn	Kaggle	7,043	Demographics, Telecom service usage, contract type, churn label	Predict customer churn & retention strategies	Pending

Case 1 – AI Driven Retail Analytics

Project Objectives

Objective	Description	Expected Output
Order Identification	Identify nature or orders	Segregation of orders into normal, return, and abnormal categories.
Automate KPI Monitoring	Automatically compute daily and monthly sales, return, and customer growth	Interactive dashboard with trend alerts
RFM Customer Segmentation	Identify customer groups by Frequency, and Monetary value	Segmentation dashboard for retention strategy
Transaction Anomaly Detection	Detect unusual purchase patterns or refund anomalies	Alert system integrated with dashboard
End-to-End Workflow Automation	Integrate data ingestion, cleaning, analysis, and reporting	Fully automated data pipeline

Case 1 – AI Driven Retail Analytics

Key achievements (Week 1)

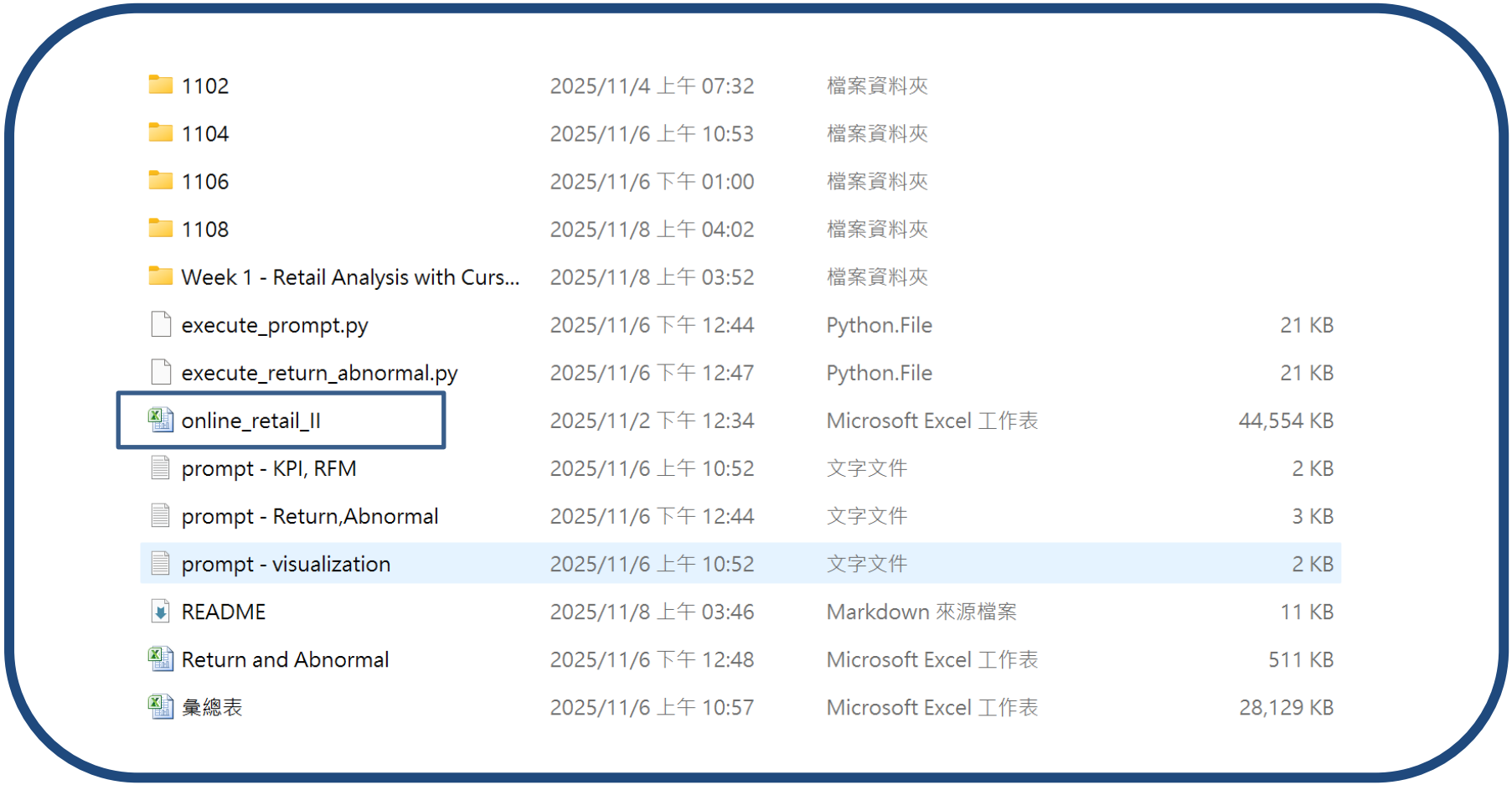
Cleaned and Identified orders into three types: normal, refund, and anomaly.
Automated **KPI monitoring** pipeline
RFM-based customer segmentation
Return Analysis and **Anomaly detection** for irregular transactions
Dynamic **reports generation** for continuous insight delivery

What's next? (Week 2 and afterwards)

Dashboard Integration
Combine KPI, Refund, RFM into unified Power BI dashboard
Advanced Analytics
Seasonality & cohort retention analysis
Revenue simulation (net vs. gross)
Deliverable
Interactive dashboard and finalize summary

Case 1 – AI Driven Retail Analytics

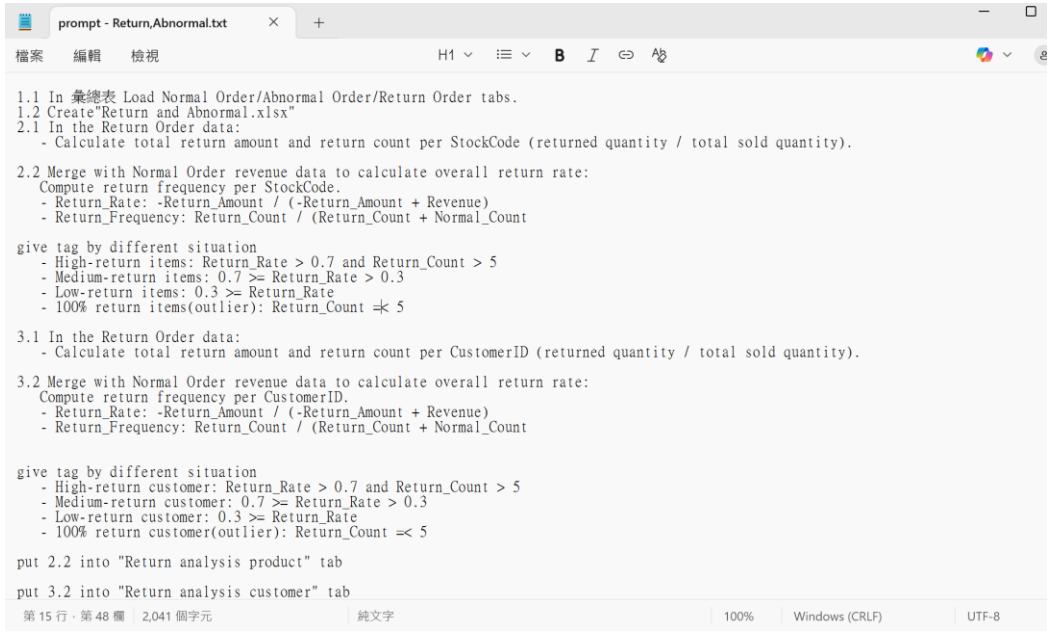
Step 1: Data Storage



1102	2025/11/4 上午 07:32	檔案資料夾	
1104	2025/11/6 上午 10:53	檔案資料夾	
1106	2025/11/6 下午 01:00	檔案資料夾	
1108	2025/11/8 上午 04:02	檔案資料夾	
Week 1 - Retail Analysis with Curs...	2025/11/8 上午 03:52	檔案資料夾	
execute_prompt.py	2025/11/6 下午 12:44	Python.File	21 KB
execute_return_abnormal.py	2025/11/6 下午 12:47	Python.File	21 KB
online_retail_II	2025/11/2 下午 12:34	Microsoft Excel 工作表	44,554 KB
prompt - KPI, RFM	2025/11/6 上午 10:52	文字文件	2 KB
prompt - Return,Abnormal	2025/11/6 下午 12:44	文字文件	3 KB
prompt - visualization	2025/11/6 上午 10:52	文字文件	2 KB
README	2025/11/8 上午 03:46	Markdown 來源檔案	11 KB
Return and Abnormal	2025/11/6 下午 12:48	Microsoft Excel 工作表	511 KB
彙總表	2025/11/6 上午 10:57	Microsoft Excel 工作表	28,129 KB

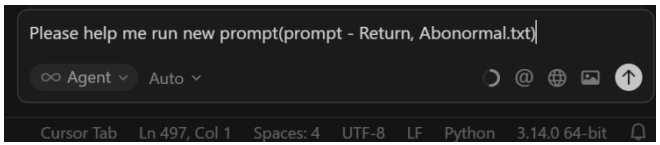
Case 1 – AI Driven Retail Analytics

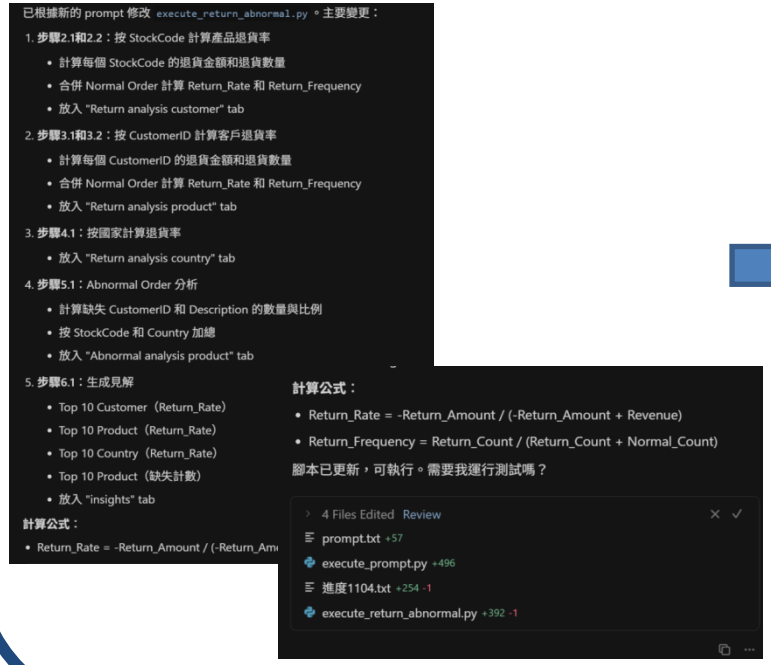
Step 2: Update prompt

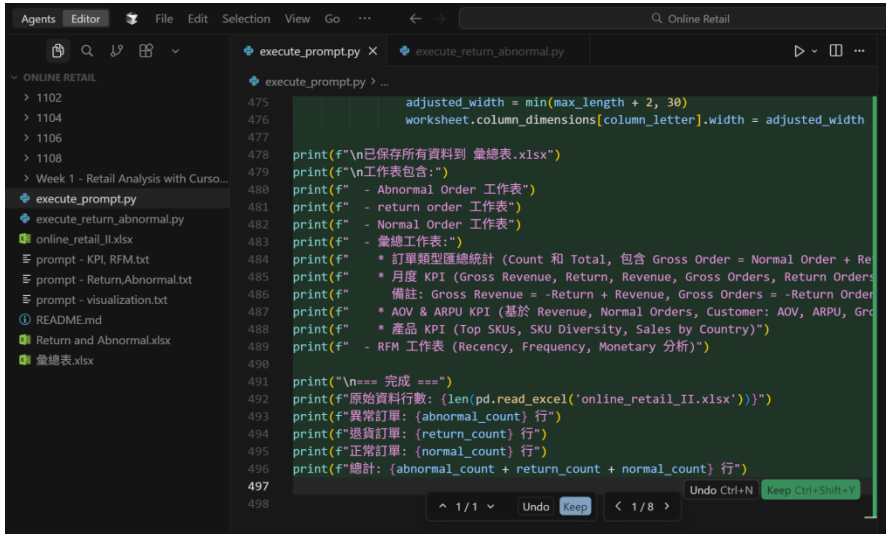


Case 1 – AI Driven Retail Analytics

Step 3: Run prompt in cursor

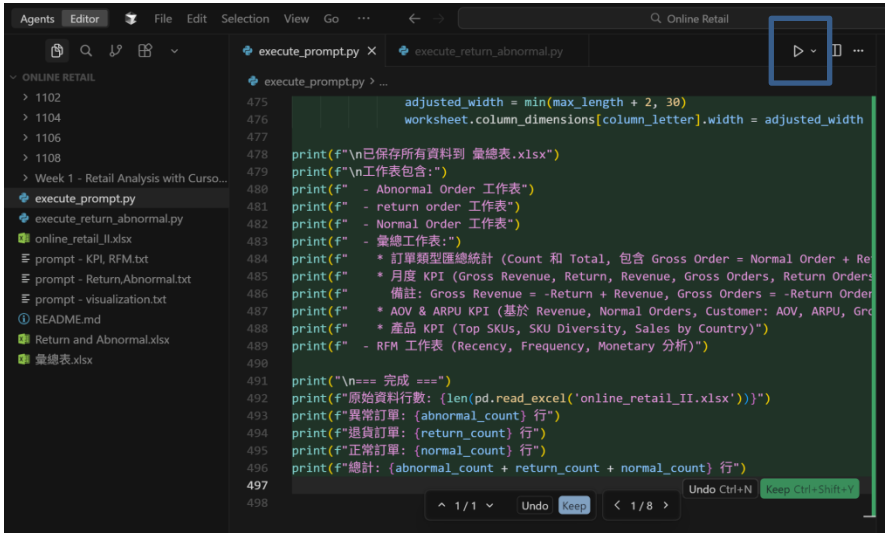


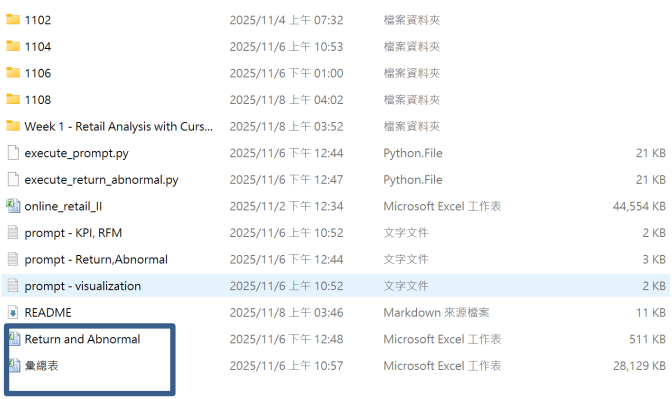




Case 1 – AI Driven Retail Analytics

Step 4: Run Script and check output





Case 1 – AI Driven Retail Analytics

Deliverable 1: Order identification

Invoice	StockCode	Description	Quantity	InvoiceDate	Price	Customer ID	Country	Total
C489449	22087	PAPER BUNTING WHITE LACE	-12	2009-12-01 10:33:00	2.95	16321	Australia	-35.4
C489449	85206A	CREAM FELT EASTER EGG BASKET	-6	2009-12-01 10:33:00	1.65	16321	Australia	-9.9
C489449	21895	POTTING SHED SOW 'N' GROW SET	-4	2009-12-01 10:33:00	4.25	16321	Australia	-17
C489449	21896	POTTING SHED TWINE	-6	2009-12-01 10:33:00	2.1	16321	Australia	-12.6
C489449	22083	PAPER CHAIN KIT RETRO SPOT	-12	2009-12-01 10:33:00	2.95	16321	Australia	-35.4
C489449	21871	SAVE THE PLANET MUG	-12	2009-12-01 10:33:00	1.25	16321	Australia	-15
C489449	84946	ANTIQUE SILVER TEA GLASS ETCHED	-12	2009-12-01 10:33:00	1.25	16321	Australia	-15
C489449	84970S	HANGING HEART ZINC T-LIGHT HOLDI	-24	2009-12-01 10:33:00	0.85	16321	Australia	-20.4
C489449	22090	PAPER BUNTING RETRO SPOTS	-12	2009-12-01 10:33:00	2.95	16321	Australia	-35.4
C489459	90200A	PURPLE SWEETHEART BRACELET	-3	2009-12-01 10:44:00	4.25	17592	United Kingdom	-12.75
C489459	90200D	PINK SWEETHEART BRACELET	-3	2009-12-01 10:44:00	4.25	17592	United Kingdom	-12.75
C489459	90200B	BLACK SWEETHEART BRACELET	-3	2009-12-01 10:44:00	4.25	17592	United Kingdom	-12.75
C489459	90200E	GREEN SWEETHEART BRACELET	-3	2009-12-01 10:44:00	4.25	17592	United Kingdom	-12.75
C489459	90200C	BLUE SWEETHEART BRACELET	-3	2009-12-01 10:44:00	4.25	17592	United Kingdom	-12.75
C489459	90185C	BLACK DIAMANTE EXPANDABLE RING	-3	2009-12-01 10:44:00	4.25	17592	United Kingdom	-12.75
C489459	90185B	AMETHYST DIAMANTE EXPANDABLE I	-3	2009-12-01 10:44:00	4.25	17592	United Kingdom	-12.75
C489459	90003C	MIDNIGHT BLUE PAIR HEART HAIR SLI	-3	2009-12-01 10:44:00	3.75	17592	United Kingdom	-11.25
C489459	90003D	CRYSTAL PAIR HEART HAIR SLIDES	-3	2009-12-01 10:44:00	3.75	17592	United Kingdom	-11.25
C489459	90003B	ROSE COLOUR PAIR HEART HAIR SLIDI	-3	2009-12-01 10:44:00	3.75	17592	United Kingdom	-11.25
C489459	90209C	PINK ENAMEL+GLASS HAIR COMB	-3	2009-12-01 10:44:00	2.1	17592	United Kingdom	-6.3
C489459	90209A	PURPLE ENAMEL+GLASS HAIR COMB	-3	2009-12-01 10:44:00	2.1	17592	United Kingdom	-6.3
C489459	90082D	DIAMANTE BOW BROOCH BLACK COL(	-2	2009-12-01 10:44:00	6.35	17592	United Kingdom	-12.7
C489476	22149	FELTCRAFT 6 FLOWER FRIENDS	-6	2009-12-01 10:55:00	2.1	13293	United Kingdom	-12.6
C489503	21540	DAIRY MAID CERAMIC BUTTER DISH	-2	2009-12-01 11:04:00	4.25	16011	United Kingdom	-8.5
C489503	21533	RETRO SPOT LARGE MILK JUG	-1	2009-12-01 11:04:00	4.95	16011	United Kingdom	-4.95
C489504	85083	KISS REINDEER SCANDINAVIAN STOCK	-6	2009-12-01 11:10:00	2.55	13916	United Kingdom	-15.3
C489518	20892	SET/3 TALL GLASS CANDLE HOLDER PI	-2	2009-12-01 11:35:00	12.75	15461	United Kingdom	-25.5
C489518	85048	15CM CHRISTMAS GLASS BALL 20 LIGI	-1	2009-12-01 11:35:00	7.95	15461	United Kingdom	-7.95
C489518	22138	BAKING SET 9 PIECE RETROSPOT	-1	2009-12-01 11:35:00	4.95	15461	United Kingdom	-4.95
C489524	21258	VICTORIAN SEWING BOX LARGE	-1	2009-12-01 11:48:00	12.75	15614	United Kingdom	12.75

Abnormal Order | Normal Order | Return Order | 匯總表

Case 1 – AI Driven Retail Analytics

Deliverable 2: KPI monitoring (Cont.)

YearMonth	Gross_Revenue	Return	Revenue	Gross_Orders	Return_Orders	Normal_Orders	Customer	Gross_Revenue_Growth	Gross_Orders_Growth	Revenue_Growth	Normal_Orders_Growth	Customer_Growth
2009-12	710036.27	-23382.11	686654.16	29762	999	30761	955	0	0	0	0	0
2010-01	582685.22	-25366.16	557319.06	21117	661	21778	720	-17.94	-29.05	-18.84	-29.2	-24.61
2010-02	523342.55	-16971.48	506371.07	22832	537	23369	774	-10.18	8.12	-9.14	7.31	7.5
2010-03	763221.5	-63612.51	699608.99	31490	812	32302	1057	45.84	37.92	38.16	38.23	36.56
2010-04	628583.36	-33974.17	594609.19	26643	595	27238	942	-17.64	-15.39	-15.01	-15.68	-10.88
2010-05	640047.03	-40061.24	599985.79	27684	960	28644	966	1.82	3.91	0.9	5.16	2.55
2010-06	706673.25	-67606.67	639066.58	30432	759	31191	1041	10.41	9.93	6.51	8.89	7.76
2010-07	620487.58	-28850.84	591636.74	26320	713	27033	928	-12.2	-13.51	-7.42	-13.33	-10.85
2010-08	621228.84	-16986.19	604242.65	25844	549	26393	911	0.12	-1.81	2.13	-2.37	-1.83
2010-09	882196.7	-50581.7	831615	33818	784	34602	1145	42.01	30.85	37.63	31.1	25.69
2010-10	1108370.22	-71690.22	1036680	48565	998	49563	1497	25.64	43.61	24.66	43.24	30.74
2010-11	1209792.8	-37456.76	1172336.04	59142	1174	60316	1607	9.15	21.78	13.09	21.7	7.35
2010-12	319132.65	-7254.65	311878	14207	298	14505	573	-73.62	-75.98	-73.4	-75.95	-64.34

YearMonth	AOV	ARPU	AOV_Growth	ARPU_Growth
2009-12	22.32	719.01	0	0
2010-01	25.59	774.05	14.65	7.65
2010-02	21.67	654.23	-15.32	-15.48
2010-03	21.66	661.88	-0.05	1.17
2010-04	21.83	631.22	0.78	-4.63
2010-05	20.95	621.1	-4.03	-1.6
2010-06	20.49	613.9	-2.2	-1.16
2010-07	21.89	637.54	6.83	3.85
2010-08	22.89	663.27	4.57	4.04
2010-09	24.03	726.3	4.98	9.5
2010-10	20.92	692.51	-12.94	-4.65
2010-11	19.44	729.52	-7.07	5.34

Case 1 – AI Driven Retail Analytics

Deliverable 2: KPI monitoring

Top SKUs

Rank	SKU	Revenue	Quantity
1	85123A	151624.31	56915
2	22423	143893.35	12502
3	M	98560.64	2631
4	85099B	85395.9	48802
5	84879	70493.83	44551
6	POST	48741.08	2212
7	21843	41718.34	4084
8	48138	41275.99	6349
9	84347	40186.65	21591
10	22086	36933.5	13860
11	20685	36431.28	5644
12	47566	35035.9	8316
13	15056N	34044.75	7201
14	85099F	33957.45	19291
15	21621	33852.96	4002
16	20914	33679.9	12356
17	20725	31933.25	20417
18	15056BL	31457.55	6709
19	22386	31149.6	17546
20	22189	29911.72	9918

SKU Diversity

YearMonth	SKU_Count
2009-12	2725
2010-01	2368
2010-02	2375
2010-03	2668
2010-04	2340
2010-05	2377
2010-06	2494
2010-07	2391
2010-08	2474
2010-09	2526
2010-10	2650
2010-11	2797
2010-12	2000

Case 1 – AI Driven Retail Analytics

Deliverable 3: FRM segregation

CustomerID	Recency	Frequency	Monetary	R_Score	F_Score	M_Score	Total_Score	Category
18245	14	13	3757.92	5	5	5	15	Champions
15298	1	12	4659.47	5	5	5	15	Champions
15299	7	6	3484.53	5	5	5	15	Champions
15301	3	6	2533.09	5	5	5	15	Champions
16150	7	6	2342.57	5	5	5	15	Champions
16161	3	25	7256.24	5	5	5	15	Champions
16168	4	40	6848.87	5	5	5	15	Champions
16186	7	15	5019.17	5	5	5	15	Champions
16191	10	11	10507.5	5	5	5	15	Champions
12431	8	13	4370.52	5	5	5	15	Champions
18061	1	19	4369.56	5	5	5	15	Champions
14680	6	43	28658.66	5	5	5	15	Champions
14560	14	22	3876.26	5	5	5	15	Champions
18144	8	8	2169.25	5	5	5	15	Champions
18168	7	11	2600.33	5	5	5	15	Champions
15271	0	10	2202.47	5	5	5	15	Champions
15281	13	7	6211.42	5	5	5	15	Champions
16125	2	14	4298.3	5	5	5	15	Champions
16126	14	11	8359.03	5	5	5	15	Champions
16131	10	6	6450.88	5	5	5	15	Champions
16133	11	23	8403.22	5	5	5	15	Champions
16143	4	7	2485.89	5	5	5	15	Champions
14085	3	16	5656.43	5	5	5	15	Champions
14092	9	10	3152.54	5	5	5	15	Champions
14099	10	9	2661.54	5	5	5	15	Champions
18075	9	10	3171.54	5	5	5	15	Champions
18087	14	15	10705.8	5	5	5	15	Champions
18092	0	8	7111.18	5	5	5	15	Champions
14702	0	13	3462.08	5	5	5	15	Champions
14709	14	12	3965.29	5	5	5	15	Champions
18102	0	89	349164.35	5	5	5	15	Champions

Abnormal Order | return order | Normal Order | 彙總 | RFM

Case 1 – AI Driven Retail Analytics

Deliverable 4: Return and abnormal analysis

Type	Rank	Detail	Return_Rate	Return_Count	Return_Amount	Missing_Total_Count	Missing_Total_Proportion
High-return items	1	D	0.9531	92	-8091.92		
High-return customer	1	15767	1	9	-242.7		
High-return customer	2	15935	1	22	-308.82		
High-return customer	3	14337	1	36	-658.63		
High-return customer	4	17485	1	6	-29.55		
High-return customer	5	17755	1	10	-72.23		
Top 10 Country by Return_Rate	1	Norway	0.4534	4			
Top 10 Country by Return_Rate	2	Malta	0.3333	2			
Top 10 Country by Return_Rate	3	Spain	0.181	43			
Top 10 Country by Return_Rate	4	Unspecified	0.1722	3			
Top 10 Country by Return_Rate	5	Korea	0.1311	10			
Top 10 Country by Return_Rate	6	Japan	0.1168	60			
Top 10 Country by Return_Rate	7	Channel Islands	0.1147	85			
Top 10 Country by Return_Rate	8	France	0.1011	240			
Top 10 Country by Return_Rate	9	Italy	0.0782	21			
Top 10 Country by Return_Rate	10	EIRE	0.0726	203			
Top 10 Product by Missing Count	1	DOT				737	100.14
Top 10 Product by Missing Count	2	22423				343	100.29
Top 10 Product by Missing Count	3	21212				326	100
Top 10 Product by Missing Count	4	47566				313	100.32
Top 10 Product by Missing Count	5	21232				307	100
Top 10 Product by Missing Count	6	85099B				271	100
Top 10 Product by Missing Count	7	22355				271	100
Top 10 Product by Missing Count	8	20724				265	100
Top 10 Product by Missing Count	9	85123A				260	100
Top 10 Product by Missing Count	10	22197				248	101.64

Return analysis product / Return analysis customer / Return analysis country / Abnormal analysis product / insights