Data Analytics Projects

By Shubha B R

Professional GitHub-ready project report containing two end-to-end projects with sample datasets and mock Power BI dashboard visuals.

Project Summary

Project 1: E-commerce Return Rate Reduction Analysis — Analyze returns, build dashboards, recommend actions.

Project 2: HR Analytics - Predict Employee Attrition — Analyze HR data, create predictive insights and dashboards.

Project 1: E-commerce Return Rate Reduction Analysis

Abstract: Analyze sample e-commerce order and return data, visualize key metrics and provide insights to reduce return rates.

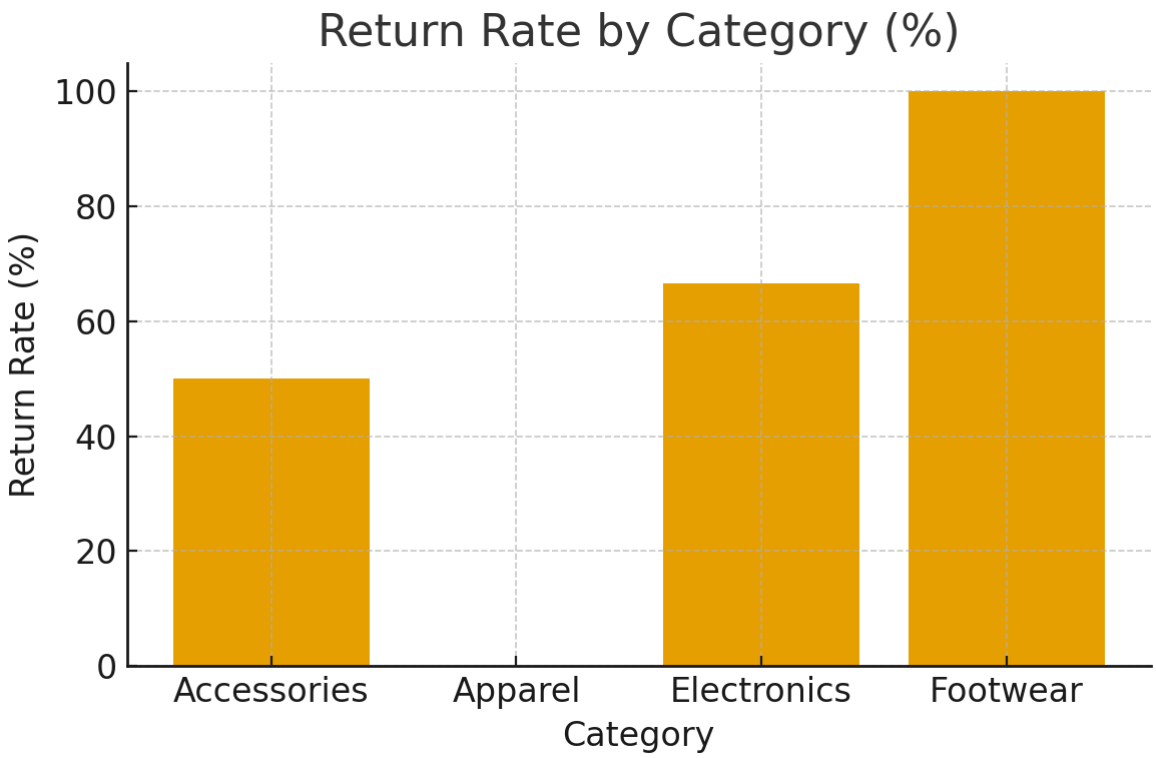
Sample Dataset (first 10 rows):

OrderID	Product	Category	ReturnReason	OrderDate	Returned
O1001	Smartphone	Electronics	Damaged	2025-07-10 00:00:00	1
O1002	T-shirt	Apparel	Size Issue	2025-07-12 00:00:00	0
O1003	Running Shoes	Footwear	Not as Described	2025-07-13 00:00:00	1
O1004	Earbuds	Electronics	Damaged	2025-07-14 00:00:00	1
O1005	Dress	Apparel	Size Issue	2025-07-15 00:00:00	0
O1006	Jacket	Apparel	Changed Mind	2025-07-16 00:00:00	0
O1007	Sneakers	Footwear	Size Issue	2025-07-17 00:00:00	1
O1008	Watch	Accessories	Damaged	2025-07-18 00:00:00	0
O1009	Backpack	Accessories	Not as Described	2025-07-19 00:00:00	1
O1010	Phone Case	Electronics	Changed Mind	2025-07-20 00:00:00	0

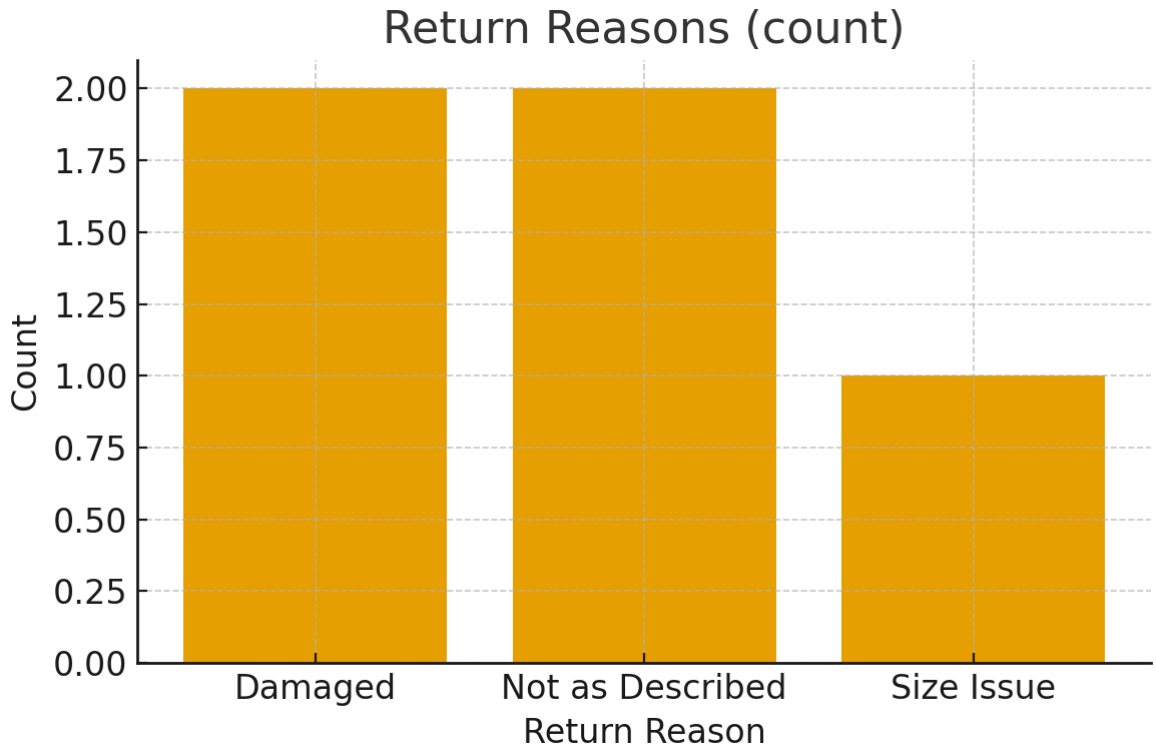
Dashboard - KPIs

Total Orders: 10 Total Returns: 5 Overall Return Rate: 50.0%

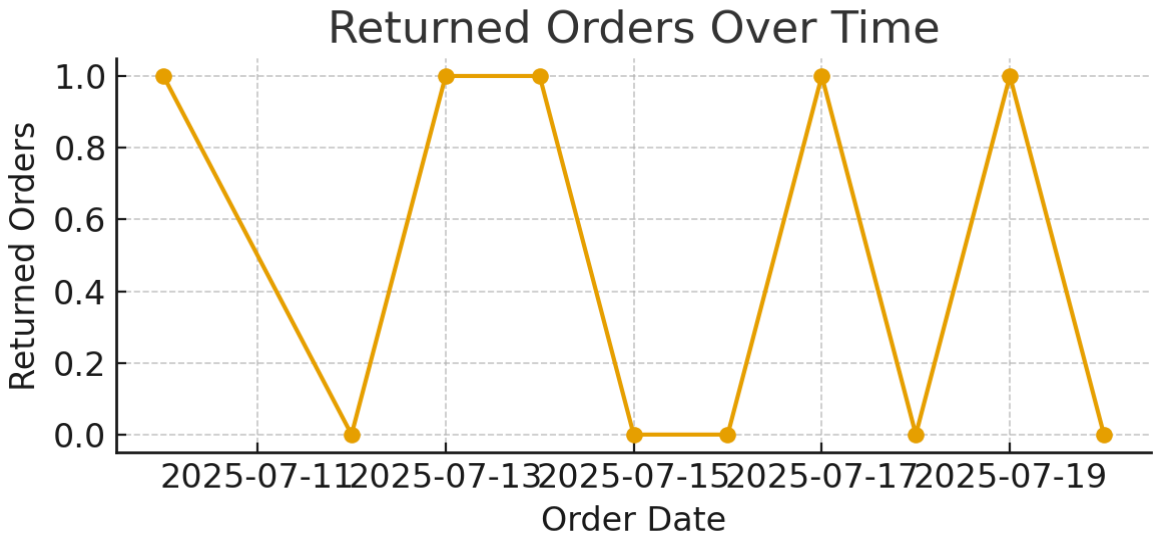
Return Rate by Category



Top Return Reasons



Return Trend Over Time


Returned Orders Over Time
Returned Orders
1.0
0.8
0.6
0.4
0.2
0.0
2025-07-11
2025-07-13
2025-07-15
2025-07-17
2025-07-19
Order Date

Insights:

Electronics and Footwear show higher return rates — focus on packaging and product descriptions.

Common return reasons: Damaged, Size Issue; improve quality checks and size charts.

Return spikes on specific dates could indicate batch issues or delivery problems.

Implementation Steps (Python + Power BI):

Use Python (Pandas) to clean and aggregate return data.

Create measures in Power BI: Return Rate = ReturnedOrders / TotalOrders.

Build visuals: KPI cards, Bar charts (by Category), Stacked bar (by Return Reason), Line chart (trend).

Publish Power BI report to Power BI Service for stakeholders.

Project 2: HR Analytics - Predict Employee Attrition

Abstract: Analyze sample HR dataset to visualize attrition drivers and provide a predictive overview.

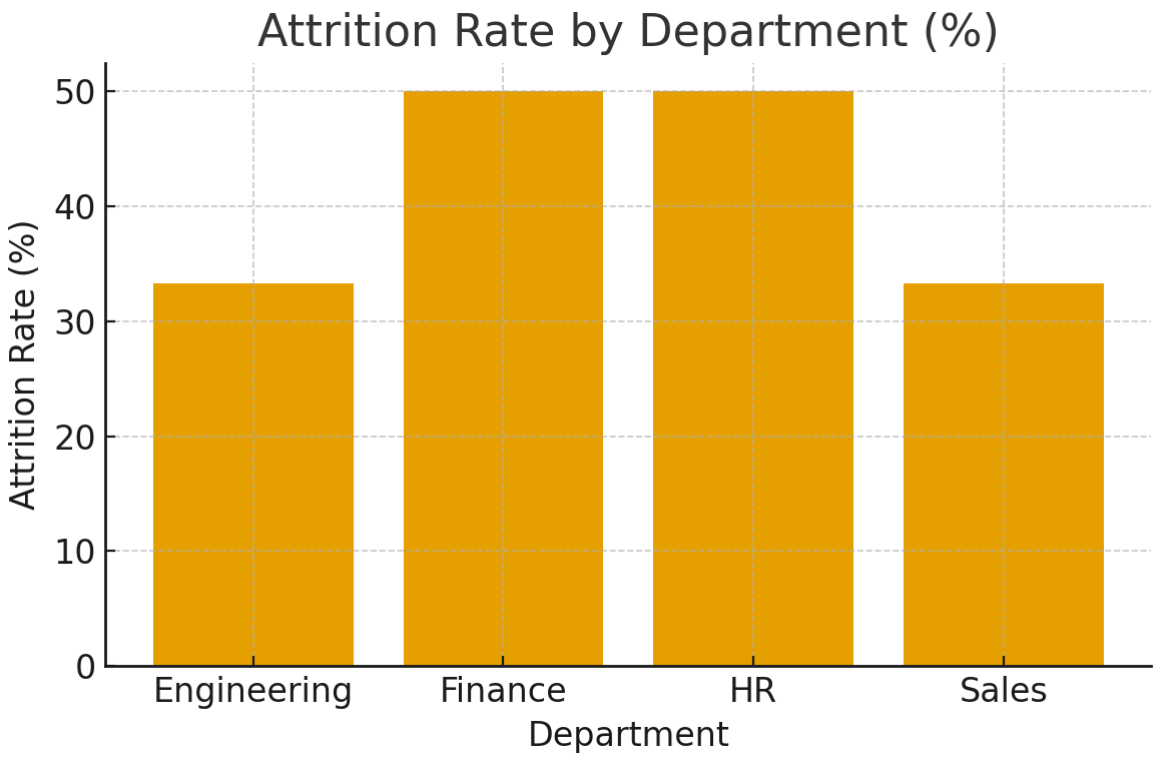
Sample Dataset (first 10 rows):

EmpID	Department	SatisfactionScore	WorkLifeBalance	PromotionLast5Y	OverTime	Age	Attrition
E001	Sales	5	3	1	2	34	0
E002	Engineering	8	4	0	3	28	0
E003	HR	3	2	1	1	40	1
E004	Engineering	6	3	0	2	30	0
E005	Sales	2	1	1	1	45	1
E006	Finance	7	4	0	3	29	0
E007	Engineering	4	2	1	2	26	1
E008	Sales	9	5	0	4	33	0
E009	HR	6	3	0	2	31	0
E010	Finance	5	3	1	2	27	1

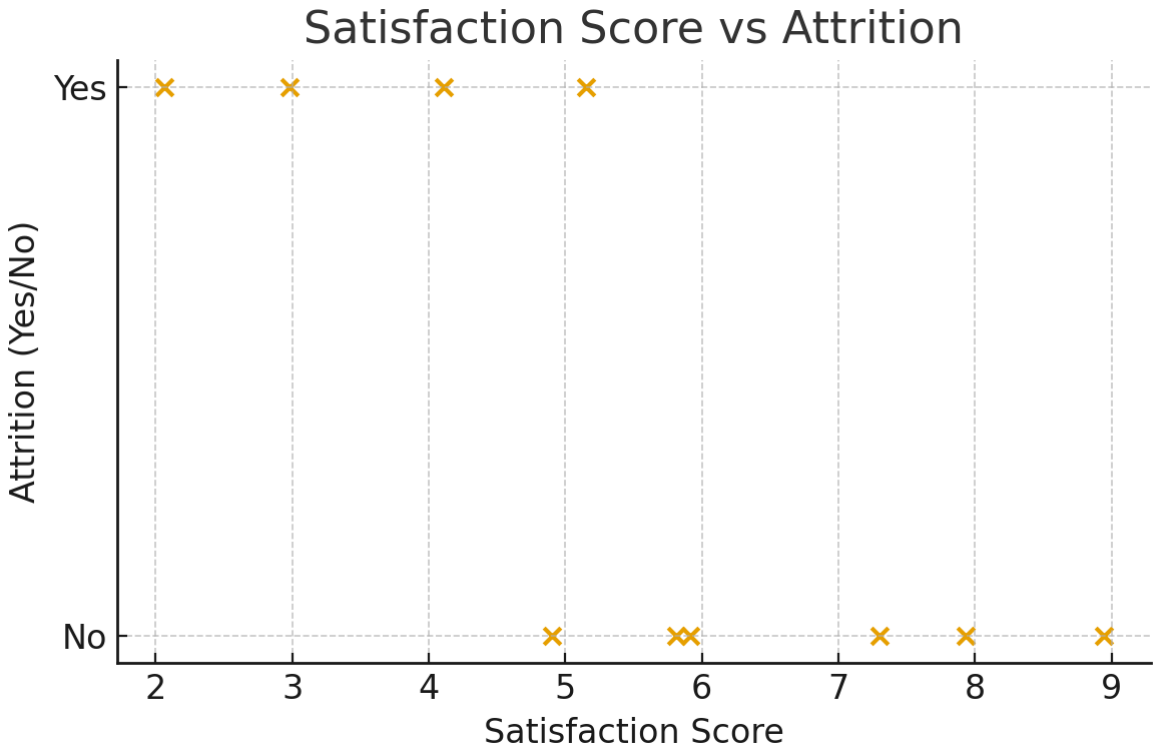
Dashboard - KPIs

Total Employees: 10 Attritions: 4 Overall Attrition Rate: 40.0%

Attrition Rate by Department



Satisfaction Score vs Attrition



Insights:

Departments with lower satisfaction scores show higher attrition — focus on engagement.

OverTime and lack of promotions correlate with higher attrition in the sample.

Use predictive models to identify at-risk employees for early intervention.

Implementation Steps (Python + Power BI + ML):

Preprocess data in Python: encode categorical variables and handle missing values.

Train classification models (Logistic Regression, Random Forest) to predict Attrition.

Evaluate with accuracy, precision, recall and confusion matrix.

Create Power BI dashboard to display predicted risk scores and recommended actions.

Final Remarks:

This combined report contains two complete project descriptions with sample datasets and mock Power BI dashboard visuals suitable for showcasing on GitHub. Use the included sample datasets to build interactive Power BI reports.