

Stage	Objective	Techniques & Tools	Expected Output
1. Data Collection	Acquire the Google Play dataset and augment with temporal/region data.	Kaggle dataset, secondary data APIs	Raw data in <code>.csv</code> format
2. Data Preprocessing	Clean, normalise, handle missing/outlier values, and create derived features.	Pandas, NumPy	Cleaned dataset
3. Exploratory Data Analysis (EDA)	Explore patterns of installs, ratings, reviews, and monetisation types.	Correlation analysis, heatmaps, pairplots, trend plots	Statistical summary and insights
4. Visualisation	Communicate findings effectively via interactive visuals.	Plotly, Seaborn, Dash, Tableau	Interactive dashboard
5. Modelling & Hypothesis Testing	Build predictive and descriptive models to answer questions.	Regression, Clustering, NLP, Classification	Predictive insights
6. Validation & Evaluation	Test model performance and interpret feature importance.	RMSE, R^2 , F1-score, Silhouette Score	Performance metrics
7. Reporting & Interpretation	Link results back to research questions and literature.	Documentation & narrative	Final report and insights