


Amazon Sales Data Analysis

Author : Deepraj Arya

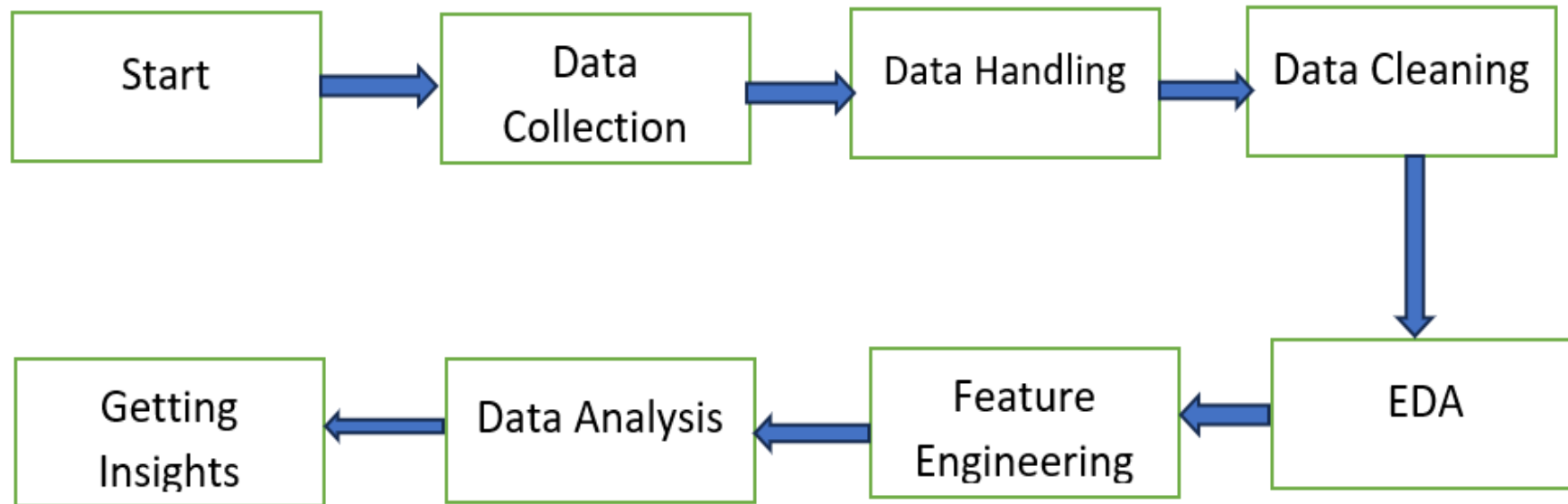
Objective:

The primary objective of this data analysis project is to gain in-depth insights into our company's sales performance and profitability, using data-driven methodologies. Through this project, we aim to achieve profitability, sales trends, key-drivers and decision making.

Benefits:

- Market insights
 - Increase revenue
 - Cost reduction
 - Decision making
- 
- A series of white lines of varying lengths and orientations are positioned in the bottom right corner of the slide, creating a modern, abstract graphic element.

Architecture

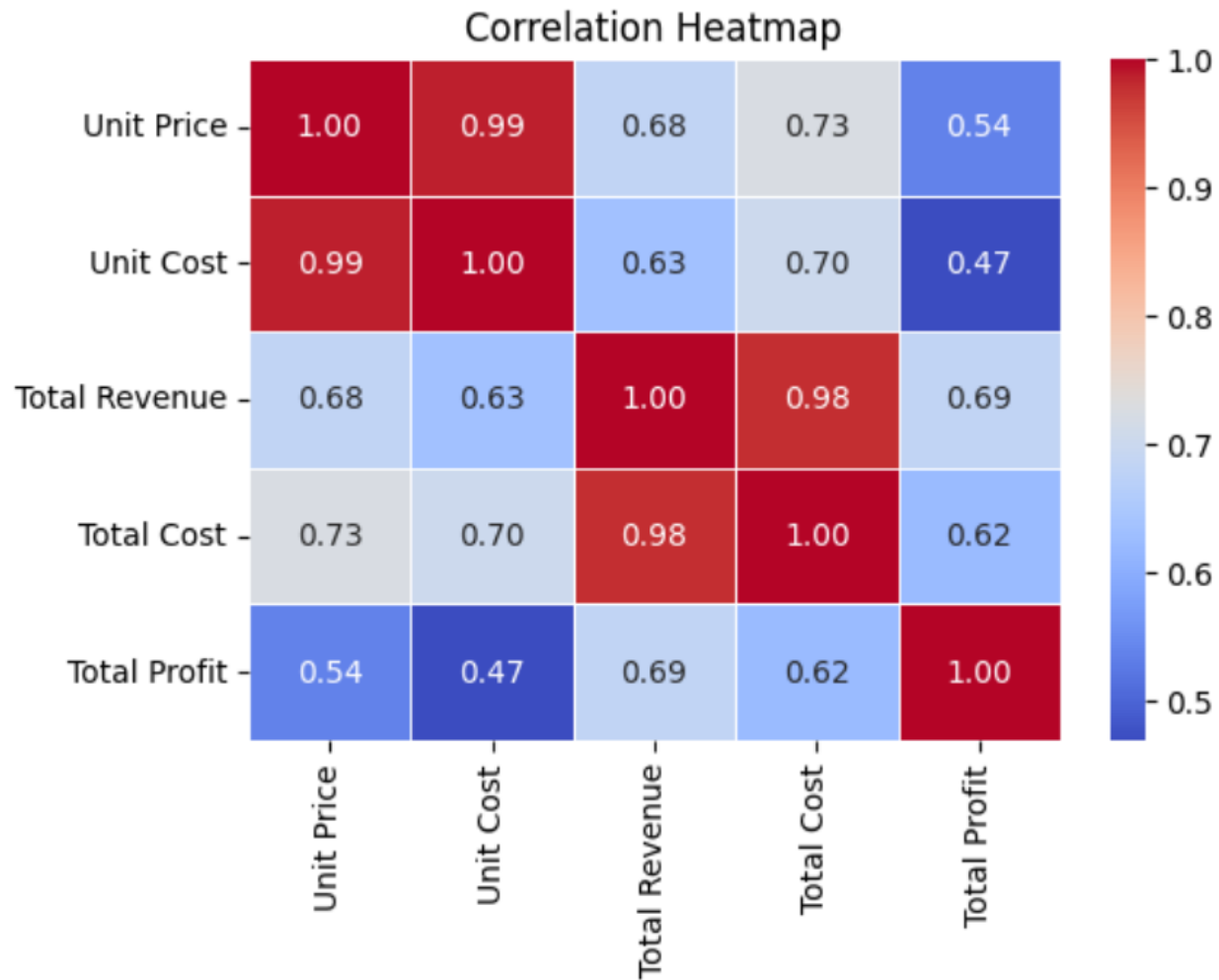


Data Pre-Processing

- ▶ The raw dataset was loaded into Pandas dataframe. The dataset was explored, organized and cleaned. The missing values were removed / replaced. New, required columns were created. The correlation between the features were checked. The cleaned dataset was loaded into Jupyter notebook and the analysis were continued.

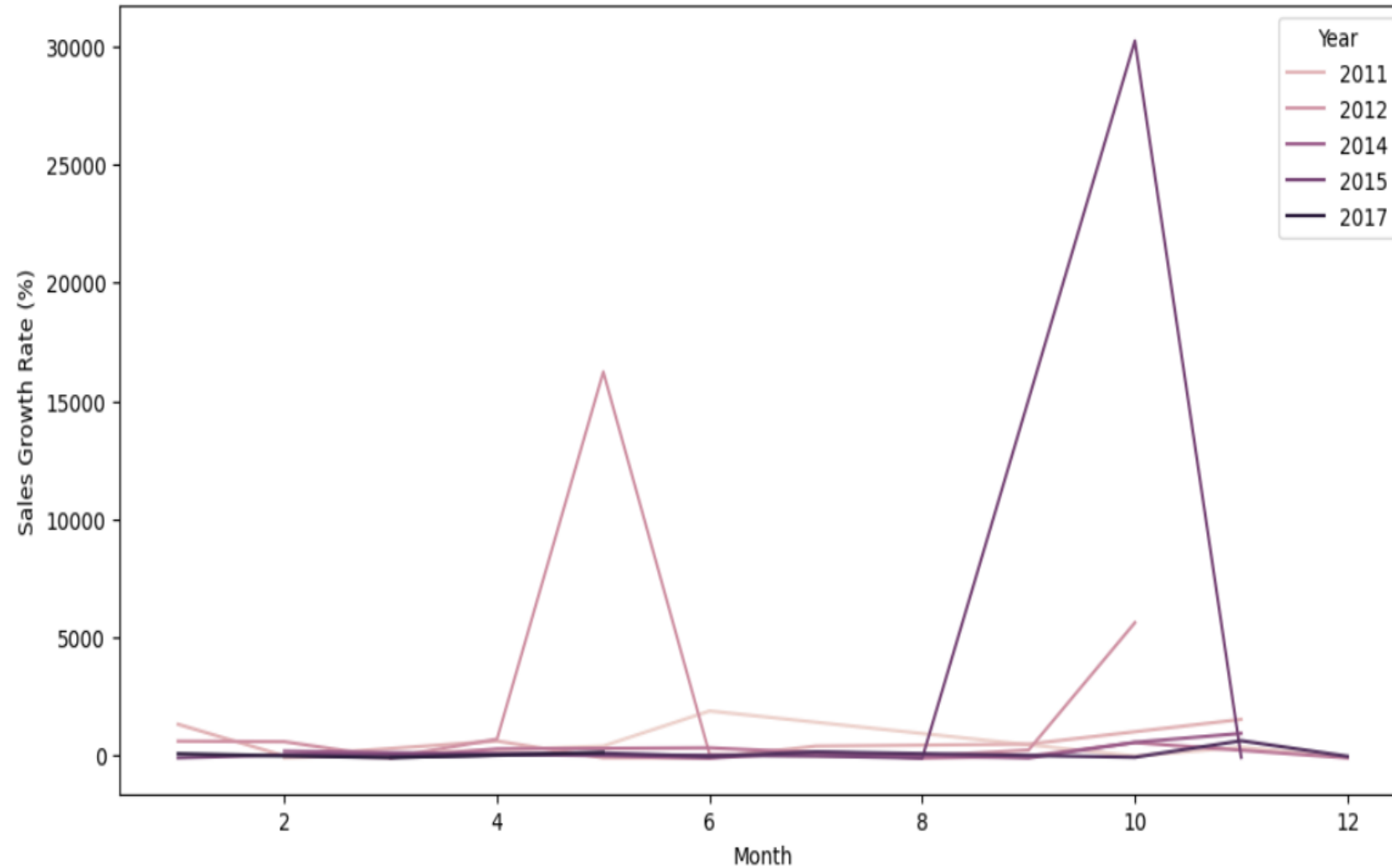


Correlation

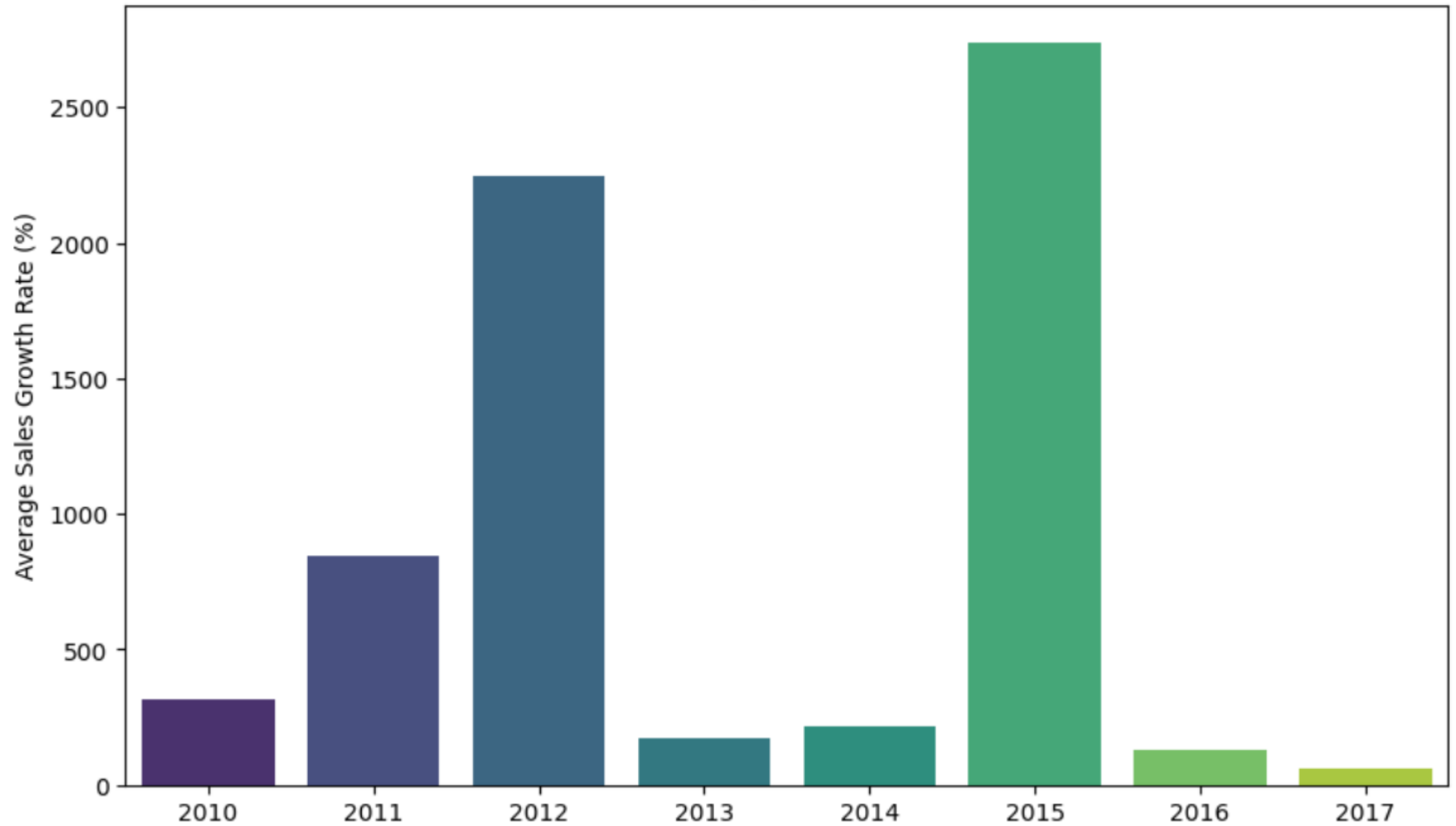


Trends


Month-Wise Sales Growth Rate Over Years



Year-Wise Average Sales Growth Rate



Insights

1. Sales increases almost triple from 2010 to 2011.
 2. Sales increases almost double from 2011 to 2012.
 3. From year 2013 to 2017 growth remains almost uniform that is around 200
 4. Year 2015 has the most growth rate > 2500
- 

Q & A:

Q1) What's the source of data?

The data for training is provided by the client in multiple batches and each batch contain multiple files

Q 2) What was the type of data?

The data was the combination of numerical and Categorical values.

Q 3) What's the complete flow you followed in this Project?

Refer slide 5th for better Understanding

Q 4) After the File validation what you do with incompatible file or files which didn't pass the validation?

Files like these are moved to the Achieve Folder and a list of these files has been shared and we removed the bad data folder.

