



Tech Sales

08/01/2022

Anthony Hayes

Data Analytics

—

Executive Summary

This dataset represents the sales information of tech from a company that sold tech such as headphones, phones and TVs. I can't

Data Sources

The data for this project is an open source data downloaded from [Sales](#)

[Analysis | Kaggle](#)

Limitation and Ethics

- The data set had the exact delivery address so had to remove the street name and number
- The states were only shown as abbreviations so had to create another data set to find the states
- The date was in American format which the UK excel translated wrong had to do some work to make it show correctly
- The data was broken down into months so used DOS to consolidate the data into 1 spreadsheet



Data Cleaning and Data Consistency Checks

- Consolidated the 12 CSV files into 1
- Removed the first line of the address
- Extracted the state from address
- Removed blank lines and checked for missing data
- Check for duplicate data
- Check for mixed data type

Data Profile

- Number of rows and columns in the cleaned dataset: 185686, 9
(Original data is, 185686, 9)

Column Details

Column	Column description	Data type	Time Variant
Order ID	The number assigned to the order	Quantitative, Discrete	no
Product	Item description	Nominal, Ordinal	no
Quantity Ordered	How many were ordered on 1 particular order number	Quantitative, Continuous	no
Date	The cost of each order	Quantitative, Continuous	no
Time	The year the boat was built	Quantitative, Discrete	Yes
Town	The length of the boat (in meter)	Nominal, Ordinal	No
Zip Code	The width of the boat (in meter)	Nominal, Ordinal	no
State	The material used to build the boat.	Nominal, Ordinal	no

Questions to Explore

- What time of day are most purchases made?
- What are the most active states?
- Is there any correlation between the value of purchases made and what day it is?