

First Glance Technical Report on a 'Sample Sales Data' for my HNG stage zero task.

(<https://hng.tech/internship>, <https://hng.tech/hire>)

Introduction: The main purpose of this report is to detect and list the anomalies from the dataset, and spot the initial insights at first glance.

Summary: From this report, it can be seen that this data consists of so many errors. The data needs thorough cleaning and preprocessing before deep analysis can be conducted. Inconsistencies and missing values hint at the possibility of other errors that can be filtered with data cleaning and validation. This data is not ready for deep analysis.

Observations:

1.) The first observation is the inconsistencies in the phone, postal code, and territories columns. Most phone numbers are entered in the wrong format and further analysis can reveal that they are incorrect. The postal code also has some wrong values entered or values entered using the wrong format. The territory column, which should have been better named 'Regions' has 'Japan' entered as a territory which is incorrect.

2.) Secondly, there are more missing values in the Addressline2 column than the inputs, so the column should be deleted entirely. The states column also has missing values for countries in the EMEA territories.

3.) The column with the name 'Territory' should be renamed 'Regions' and there should be a data dictionary to help understand the meaning of the column with the name 'MSRP'.

- 4.) There is no currency symbol on the 'Price_Each' column.
- 5.) The data also shows that values were recorded in the sales column for canceled orders which doesn't seem right.
- 6.) One of the quickest insights from this data is, that there is no relationship between sales and years. This shows that sales neither increased nor decreased as the year progressed.

Conclusion: In conclusion, this dataset can be said to contain many errors and needs preprocessing before deep analysis is conducted on the dataset to avoid drawing false insights from the data and making wrong business decisions.