

Prompt: Outlier Detection

Instruction

- Tutorials: [Prompt Engineering for Outlier Detection - KDnuggets](#)
- ???????????? LLM ?????????????????????????????????

Prompt to Detect Outliers:

You are a data analysis assistant. I have attached a dataset. Your task is to detect outliers using three methods: Standard Deviation, IQR, and Percentile.

Follow these steps:

1. Load the attached dataset and remove both the "\$" sign and any comma separators (",") from financial columns, then convert them to numeric.

2. Handle missing values by removing rows with NA in the numeric columns we analyze.

3. Apply the three methods to the financial columns:

Standard Deviation Method: flag values outside $\text{mean} \pm 3 * \text{std}$

IQR Method: flag values outside $Q1 - 1.5 * IQR$ and $Q3 + 1.5 * IQR$

Percentile Method: use the 1st and 99th percentiles as cutoffs

4. Instead of listing all results for each column, compute and output only:

- the total number of outliers detected across all financial columns for each method
- the average number of outliers per column for each method

Additionally, save the row indices of the detected outliers into three separate CSV files:

- sd_outlier_indices.csv
- iqr_outlier_indices.csv
- percentile_outlier_indices.csv

Output only the summary counts and save the indices to CSV.

```
financial_columns = [  
"ipa_funding",  
"ma_premium",
```

```
"ma_risk_score",  
"mbr_with_rx_rebates",  
"partd_premium",  
"pcp_cap",  
"pcp_ffs",  
"plan_premium",  
"prof",  
"reinsurance",  
"risk_score_partd",  
"rx",  
"rx_rebates",  
"rx_with_rebates",  
"rx_without_rebates",  
"spec_cap"  
]
```

Prompt to Remove the Outliers:

You are a data analysis assistant. I have attached a dataset along with a CSV which includes indices which are outliers.

Your task is to remove these outliers and return a clean version of the dataset.

1. Load the dataset.
2. Remove all given outliers using the given indices.
3. Confirm how many values were removed.
4. Return the cleaned dataset.

Revision #2

Created 2025-12-10 10:38:14 CST by A-Lang (Admin)

Updated 2025-12-10 10:49:15 CST by A-Lang (Admin)