**Data Analysis**

**Part 1:**

The CFO of your firm has requested a report to help analyze potential fraudulent transactions. Using your newly created database, generate queries that will discover the information needed to answer the following questions, then use your repository's ReadME file to create a markdown report you can share with the CFO:

* Some fraudsters hack a credit card by making several small transactions (generally less than $2.00), which are typically ignored by cardholders.
  + How can you isolate (or group) the transactions of each cardholder?
  + Count the transactions that are less than $2.00 per cardholder.
  + Is there any evidence to suggest that a credit card has been hacked? Explain your rationale.
* Take your investigation a step futher by considering the time period in which potentially fraudulent transactions are made.
  + What are the top 100 highest transactions made between 7:00 am and 9:00 am?
  + Do you see any anomalous transactions that could be fraudulent?
  + Is there a higher number of fraudulent transactions made during this time frame versus the rest of the day?
  + If you answered yes to the previous question, explain why you think there might be fraudulent transactions during this time frame.
* What are the top 5 merchants prone to being hacked using small transactions?
* Create a view for each of your queries.

#### Part 2:

Your CFO has also requested detailed trends data on specific card holders. Use the [starter notebook](http://localhost:8888/files/ColumbiaFintechLab/cu-nyc-virt-fin-pt-03-2021-u-c/02-Homework/07-SQL/Instructions/Starter_Files/challenge.ipynb?_xsrf=2%7Cc53c707f%7Cb08b55cc237c1656b52eb12cf0cad7f6%7C1618081049) to query your database and generate visualizations that supply the requested information as follows, then add your visualizations and observations to your markdown report:

* The two most important customers of the firm may have been hacked. Verify if there are any fraudulent transactions in their history. For privacy reasons, you only know that their cardholder IDs are 2 and 18.
  + Using hvPlot, create a line plot representing the time series of transactions over the course of the year for each cardholder separately.
  + Next, to better compare their patterns, create a single line plot that contains both card holders' trend data.
  + What difference do you observe between the consumption patterns? Does the difference suggest a fraudulent transaction? Explain your rationale.
* The CEO of the biggest customer of the firm suspects that someone has used her corporate credit card without authorization in the first quarter of 2018 to pay quite expensive restaurant bills. Again, for privacy reasons, you know only that the cardholder ID in question is 25.
  + Using Plotly Express, create a box plot, representing the expenditure data from January 2018 to June 2018 for cardholder ID 25.
  + Are there any outliers for cardholder ID 25? How many outliers are there per month?
  + Do you notice any anomalies? Describe your observations and conclusions.