Analyst Project

Data

Attached in the project file will be three CSVs containing records for Opportunities, Accounts, and Sales Representatives. Below is a description of the data and how the tables relate to one another.

OPP\_ID: Unique key for Opportunity records.

ACCOUNT\_ID: ID representing the Account we are trying to sell to. OWNER\_REP\_ID: ID representing the Sales Representative that owns that Opportunity.

OUTCOME: The outcome of the Opportunity – ‘Closed Won’ or ‘Lost’. CREATED\_DATE: The date the Opportunity was created – Opportunities are in the first sales stage (‘Discovery’) when created.

CLOSE\_DATE: The date the Opportunity was closed.

DATE\_OF\_QUALIFYING: A timestamp of the most recent time an Opportunity was moved into the second sales stage – ‘Qualifying’.

DATE\_OF\_EVALUATION: A timestamp of the most recent time an Opportunity was moved into our third sales stage – ‘Evaluation’.

DATE\_OF\_PROCUREMENT\_NEGOTIATIONS: A timestamp of the most recent time an Opportunity was moved into our fourth sales stage – ‘Procurement/Negotiations’. DATE\_OF\_VERBAL\_PENDING: A timestamp of the most recent time an Opportunity was moved into our fifth sales stage – ‘Verbal Pending Close’.

**Sales Rep Table:**

Each record in this table represents an individual responsible for selling PitchBook Subscriptions.

REP\_ID: Unique key for Sales Representative records.

OFFICE: The regional PitchBook office the Sales Representative works in. NAME: The name of the Sales Representative.

EMAIL: The email address of the Sales Representative.

**Account Table:**

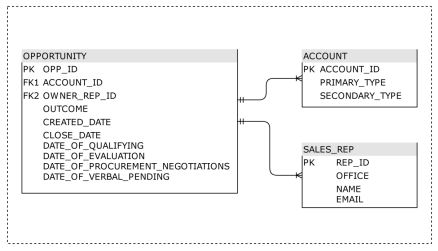
Each record in this table represents an Account in our sales system.

ACCOUNT\_ID: Unique key for Account Records.

PRIMARY\_TYPE: The primary type of the Account.

SECONDARY\_TYPE: The secondary type of the Account.

**Entity Relationship Diagram – how the tables relate:**

****

Questions for Analysis

Please use the data provided to effectively analyze and communicate your answers to the questions below. Feel free to use any data manipulation, analysis, and visualization tools that you are comfortable with. We encourage you to create any data visualizations that would support and communicate your responses – bear in mind we are also evaluating your analytical and data communication skills. Finally, clearly document your work and submit all associated files - SQL, Python, Excel Workbooks, Tableau Workbooks, Power BI Report etc.

1. What is the mean and median number of days to close (the difference in days between “CREATED\_DATE” and

“CLOSE\_DATE”)?

a. In what situations do you think each measure is more relevant?

b. Do they meaningfully differ overall?

c. Do they differ by Outcome?

d. Do they differ by Office?

e. Do they differ by Primary Type?

2. What is the mean and median duration of Opportunities in the Evaluation stage?

a. Does this differ by Primary Type?

b. Does the duration of an Opportunity in the Evaluation stage correlate with a given Outcome?

3. From your answers to the questions above, what additional insights do you have about our business?

Assumptions:

1. If there was no Evaluation date I excluded it from the data since technically it had never started the evaluation stage.
2. If there was no date in the next stage (procurement/negotiation) or (Verbal/pending)I used the Closed date as the date to subtract

Tools Used:

1. Excel - for data cleaning/manipulation
2. MySQL - for data cleaning/analyzing
3. Tableau - for Visualization/analyzing