Guidelines and Expectations

- Complete this assignment prior to your in-person interview.
- Be prepared to give a 20-30 minute presentation on your findings. Cater your presentation towards a leadership level audience. They will understand the high-level goals of the project but not necessarily know the details.
- Submit your presentation and answers in Google slides or PowerPoint.
- Email completed presentation to the recruiter you are in contact with.
- Pivot Table usage encouraged but any other data analysis/manipulation techniques are certainly allowed.
- If you have an exciting data modeling skill you want to show off (R/Python), here is your chance!
- Please read through the entire assignment before starting your work.
- Work independently. No collaboration allowed.

Intro:

You are a member of the athenahealth product team, supporting several development teams working on exciting new functionality. Three months ago, your team released a new product feature to a set of test clients. Today, you have a data set that contains measurements from the first few months of the product test. Your job is to review this raw data to get a better understanding of how clients are using the new functionality. Ideally, the insight you glean from this data analysis will help you to better inform your teams of future development work.

The Feature

One of the main goals at athena is to make it easy for our clients to get paid for the patient related services they provide. Unfortunately, there is still a component of manual work that is required of our clients; whether it's adding a medical billing code to a claim for the patient services rendered or appealing a claim for a lower than expected payment.

The new product feature that your team has released was designed to improve these currently existing workflows (listed above). However, despite having access to the new product feature, clients are still able to take an "off ramp" to utilize the old legacy functionality for the same existing workflows. Specifically, there are two classes of actions that can be taken by clients: *Coding*, and *Appeal* actions. These actions can be taken by clients utilizing the new product feature, as well as by using the old legacy functionality. The accompanying data set highlights the number of times users are taking *Coding* and *Appeal* actions using either the new product feature or the old legacy functionality.

Assignment:

- **In 2-4 powerpoint/google slides, summarize and group the overall data to tell the story of how clients are using the new feature (if at all).**
  Hint: The data set is designed to be ambiguous and the answers to questions open ended. You will have to make your own assumptions about how to structure the data and what to exclude or include. Please make any assumptions you make clear in your presentation.

    - What % of clients are using this new functionality?
    - What types and sizes of clients are using the functionality the least/the most?
    - Do larger clients tend to correlate with higher usage rates?

- **Spend 4-6 slides to answer the following questions:**
    - Create 2 metrics (with visualizations) that can be utilized to track the success of this new functionality. Why did you choose these metrics?
    - Has this new functionality proven to be successful? Are there any significant trends?
    - Would you recommend that this functionality gets expanded to the entire client base? Or are there certain clients that should be expanded to first?
    - We have access to a product education team at athena that creates training materials and walkthroughs for our clients. Where would you recommend they focus their efforts?
    - Are there any pieces of information that would be useful to have in the next version of this data set?

Data:

Attached to the email is a data set titled ‘athena_interview_data.csv’. This data set is entirely made up, but represents the real data and questions we strive to answer. The data you have been given has two sheets. *Hint: Sheets can be combined*

1) **Claim Actions Data**

This sheet contains data about how the claim action page has been utilized over the past 3 months. Each row represents 1 week for each client. The data includes columns describing:

- CLIENTID: ID to identify each client
- WEEK: Time period in which the claim actions occurred
- LEGACYCODINGACTIONS: Number of coding actions performed by clients using the legacy functionality
- NEWCODINGACTIONS: Number of coding actions performed by clients using the new product functionality
  LEGACYAPPEALACTIONS: Number of appeal actions performed by clients using the legacy functionality
- NEWAPPEALACTIONS: Number of appeal actions performed by clients using the new product functionality

2) **Client Details Data**

This sheet contains basic demographic information about our clients. You might find it helpful in your analysis.

- CLIENTID: ID to identify each client
- OVERALLSTATUS: Client contract status
- CLIENTSPECIALTY: Client’s medical specialty
- CLIENTSIZE: General size of each client