Practical Exam - Healthcare revenue

From:	Head of Data Science
Received:	Today
Subject:	New project from the Financial Planning Team

Hey!

I have a new project for you from the Financial Planning Team. It should be an interesting challenge. You can see the background and request in the email below.

I would like you to perform the analysis and write a short report for me. I want to be able to review your code as well as read your thought process for each step. I also want you to prepare and deliver the presentation for the planning team - you are ready for the challenge!

The team need to better understand the drivers of revenue to help the growth of the medical group. They also want us to predict revenue within 10% of the actual values. I am not sure if that is realistic in the time we have, but do your best and present whatever you find.

You can find more details about what I expect you to do here. And information on the data here.

I will be on vacation for the next couple of weeks, but I know you can do this without my support. If you need to make any decisions, include them in your work and I will review them when I am back.

Good Luck!

From:	Senior Financial Analyst
To:	Head of Data Science
Received:	Yesterday
Subject:	Can you help us predict revenue?

Hi,

We have a new project that will have a high impact to the medical group and I am hoping your team can help us to build something to help.

In brief, we need to optimize revenue generation while maintaining high-quality patient care. We know that various factors could be impacting the revenue, but not to what extent. So that we can plan effectively, we need to understand what drives revenue. Ideally we would also like to be able to predict how changes will impact revenue.

Can your team help us to:

- Gain a better understanding of the main variables influencing revenue generation?
- Identify whether we can predict within 10% of the actual revenue values to enable us to predict how changes might impact revenue?

We are currently making plans for the next financial year and this information is vital, so I need you to present your results to me by the end of the month. Whatever your results, what do you recommend we do next?

I look forward to seeing your presentation.

About FMPF Medical Group

FMPF Medical Group, with its diverse array of specialized and general hospitals and clinics, stands at the forefront of healthcare.

What once started as a small mental health institute has blossomed into a vast conglomerate comprising numerous specialized and general hospitals and clinics.

FMPF Medical Group now proudly serves a wide range of patients with diverse medical needs and has consistently pursued improvement over the years, demonstrating unwavering dedication to delivering comprehensive and compassionate care.

Collecting Clinic Data

The hospitals and clinics in the FMPF Medical Group are located across the country in a mix of urban (cities or large towns) and rural (small towns and villages) locations.

Some of the locations are expected to be more profitable, and tend to offer specialist treatment to private, paying patients. Other locations are not expected to make a profit and offer general care to anyone who needs it.

The clinical data provided, including patient numbers and ratios, are provided by each of the locations. Each clinic is asked to provide the data for a random day in the year. The intention is to get a snapshot of the particular clinic.

While some locations have tracking systems to automatically calculate the numbers based on daily events, other locations have to track the information manually.

Satisfaction scores are obtained from patients who complete a short survey provided to them as they are leaving. The questions given to each patient are shown below.

1. How long did you have to wait on this visit before seeing a medical professional?
2. How long (in hours) did you spend on our site in total on this visit?
3. Were you able to get the attention of a nurse when you needed one?
4. Were you happy with the care you received on this visit? (Score 1-10)

Data Information

The data is collected from various clinics and hospitals across the country over the past year. The data does not correspond to a single date, but a random date across the year for each clinic.

You can find the data here, as well as inside the Workspace template for this task.

I will let you decide how to process it, just make sure you include all your decisions in your report.

Don't forget to double check the data really does match what they say - it might not.

Column Name	Details
clinic_id	Discrete. Unique identifier for the clinic.
ownership	Categorical. The ownership type of each medical facility (possible values: Profit, Non-Profit, Public).
hospital_location	Categorical. The location type of each medical facility (Urban, Rural).
admitted_patients	Discrete. The number of patients admitted to each medical facility on that day. Any positive integer.
nurse_patient_ratio	Continuous. The ratio of nurses to patients (values ranging from 0.0 to 1.0).
ed_visits	Discrete. The number of emergency department visits during that day. Any positive integer.
satisfaction_score	Continuous. The satisfaction score given by patients about their experience (1-10).
revenue	Continuous. Daily revenue value. Any positive value to two decimal places.

Guide to Data Science Projects

1. I would like you to create a written report to summarize the analysis you have performed and your findings. The report will be read by me (Head of Data Science). The list below describes what I expect to see in your written report.

2. You will need to use DataCamp Workspace to complete your analysis, write up your findings and share visualizations.

3. You must use the data provided for the analysis.

4. You will also need to prepare and deliver a presentation. You should prepare around 6-10 slides to present. The list below describes what they expect to see in your presentation.

5. Your presentation should be no longer than 10 minutes.

Written Report

Your written report should include written text summaries and graphics of the following:

- Data validation:
  - Describe validation and cleaning steps for every column in the data
- Exploratory Analysis to answer the customer questions ensuring you include:
  - Two different types of graphic showing single variables only
  - At least one graphic showing two or more variables
  - Description of your findings
- Model Development including:
  - What type of problem this is
  - Fitting a baseline model
  - Fitting a comparison model
- Model evaluation
  - Show how the two models compare
- Definition of a metric for the business to monitor
  - How should the business monitor what they want to achieve?
  - Estimate the initial value(s) for the metric based on the current data?
- Final summary including recommendations that the business should undertake

Presentation

You will give an overview presentation to the individual who requested the work. The presentation should include:

- An overview of the project and business goals
- A summary of the work you undertook and how this addresses the problem
- Your key findings including the metric to monitor and current estimation
- Your recommendations to the business

Grading

Before submitting your written report or delivering your presentation, remember to check your work against the grading criteria.

You can find the full grading rubric here.

You must pass all criteria to pass this part of the certification.