Sample Practical Exam - Revenue at Solstice Opal Hotel

You are a Data Analyst who has just joined Solstice Opal Hotel, a luxury hotel. You have received the following email from your manager:

From:	Head of Analytics
Received:	Today
Subject:	New task from leadership

Hi,

Welcome again to Solstice Opal Hotel — I'm glad you're here. I want to get you started on something that's top of mind for our leadership team.

Over the past quarter, we've seen big differences in ancillary revenue (spa, dining, and activities) across our guest base. Some guests barely use any of our services, while others spend a lot — but right now we don't have a clear view of which types of guests are driving that revenue or where we might be leaving money on the table.

Leadership has asked us to come back with a clear picture of which guest segments generate the most ancillary revenue, and practical, data-informed actions to maximise this revenue source.

To help you get started, we've pulled together a dataset from three different systems (Guest profiles, Stay details, Ancillary spend). The data is not perfect; this is the reality of how our data comes together today — different teams, different systems, and not always consistent definitions. Part of the value you'll bring is making sense of this.

I would like you to perform the analysis and write a short report for me. I don't need to see any code, but I do want to read your thinking, how you handled any data cleaning and interpretation, and how you got to your conclusions.

I also want you to prepare and deliver the presentation to senior leadership. Remember they're not data specialists and will be focussed on what they can do to maximise revenue.

I've already asked the Revenue Manager and Engineering Lead a few questions about how this data is collected and used; you can see their answers below.

You can find more details about what I expect you to do 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.

Thanks!

From:	Revenue Manager
Received:	Today
Subject:	Product and data clarification

Hi,

Happy to share a bit more context on the product and the data. Most of it is fairly straightforward, but there are a few things worth calling out:

Stay details:

- check_in_date is the actual date the guest arrived.
- booking_channel should show where the booking originally came from.
- number_of_guests reflects how many people actually stayed, not how many were on the initial reservation.
- reason_for_stay: our guests tend to be either business travellers or leisure travellers, and they usually stick to one or the other — it's pretty rare for someone to switch between the two.
- Guests shouldn't have overlapping stays; people generally don't book multiple rooms for themselves.

Guest profiles:

- The loyalty tier you see is whatever tier the guest held at the time of their stay, not their current one.
- Marketing_consent indicates whether a guest has agreed to receive promotional communication

Ancillary spend:

- This dataset is fairly simple, it records ancillary purchases guests made.
- All ancillary categories have roughly the same margin, so we don't treat any one category as more strategically important than another.

I'm not completely sure how often the data refreshes — the Engineering Manager may know more. The data you've got should give you a good baseline; we don't see a lot of seasonal variation in our location.

For your recommendations, the most valuable ones will be clear, actionable steps that leadership can take immediately, backed up by solid evidence. They're not keen on overly technical suggestions — they want to know what to do next to boost results.

Best,
Chris (Revenue Manager)

From:	Lead Engineer
Received:	Today
Subject:	Data definition and pipelines

Hey there,

Happy to clear a few things up about the data.

The three datasets all come from different source systems, built separately by different teams over the years. There isn't really a shared set of standards or checks across them, so it's pretty normal to see missing fields or things that don't line up perfectly—plus our systems do have the occasional hiccup.

We don't do much processing in the pipelines. Apart from basic structuring to make the data ingestible, there's no real cleaning or standardizing happening before the data lands with you.

Everything updates once a day.

Marketing consent lives in a separate permissions service we integrate with. For the datasets you're working with, consent flags are stored for reference but aren't used to filter or mask any records — nothing is dropped or anonymised based on consent status at this stage of the pipeline.

Loyalty tier reflects whatever the guest's tier was at the time of their stay, not their current level.

In Ancilliary_spend, each row is a single transaction tied to a guest_id. It won't always map neatly to a specific stay so I wouldn't worry about that too much.

Hope that helps!

Sasha (Lead Engineer)

Data Information

The data hasn't been validated, so make sure that you check it against all of the information in the table before you start your analysis.

Download the data by clicking these links:
[Guest profiles]
[Stay details]
[Ancillary spend]

DataFrame	Columns
guest_profiles	guest_id, loyalty_tier, marketing_consent
stay_details	guest_id, check_in_date, stay_id, booking_channel, reason_for_stay, number_of_guests
ancillary_spend	Guest_id, category, amount_spent

Guide to Analysis 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 Analysis). The list below describes what I expect to see in your written report.

2. You will need to use a DataLab workbook to 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 8-10 slides to present to senior leadership. 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 questions I've given you 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
- 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 senior leadership. 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.