Sample Practical Exam - Used Car Sales

We have included a [sample solution](#) at the end of this document. Check it out to see what a passing solution to this project looks like.

From:	Head of Data Science
Received:	Today
Subject:	New task from the sales team

Hey!

I have an analysis task for you from the sales team. You can see the background and request in the email below. They were quite skeptical about how we could help them, so this is a great opportunity to start to improve processes and make the team more efficient.

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 sales team - you are ready for the challenge!

They want us to predict prices within 10% of the listed price. But as their team can only manage 30%, it is probably ok to show we are at least as good as that. I don't know how close you will get 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 in the "Guide to Data Science Projects", included below. And information on the data in the section "Data Information".

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:	Sales Team
To:	Head of Data Science
Received:	Yesterday
Subject:	Selling Price

Hi,

It was great to meet you and the team last week. I have been thinking and there might be a project you could help us out with.

Next month our most experienced sales team member will be retiring. They have been on the team almost since the company was founded. They are incredibly talented at estimating the sales price of cars. We are quite worried that when they retire we won't be able to estimate as well and that will have a huge impact on sales.

Currently, when a new car comes in, team members take all of the information that usually appears in the advert and give it to this team member. They then estimate the price. We have been testing the team members estimating themselves but they are always around 30% away from the price we know the car will sell for.

Can you help us estimate the price we should list a car for? The team estimates are always around 30% off, we really want to be within 10% of the price. This will mean we can automate the whole process and be able to sell cars quicker.

As I said, the team member retires in a month, so we would like to get your initial thoughts as soon as possible. We would like to see a presentation, you will be presenting to me and another sales manager. We would like to hear about whatever you manage to achieve to help us make decisions on the way forward.

Look forward to seeing your presentation.
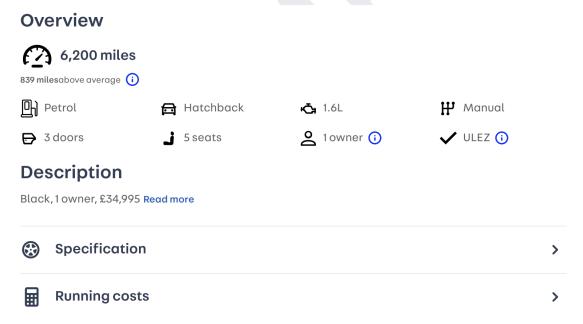
Company Background

Discount Motors is a used car dealership in the UK. They want to lead the way in used cars. Selling to customers who want the latest and greatest features, without the price tag of a brand new car.

The UK Government has now announced that from 2030 all new cars will be required to be zero emissions. Although this won't impact the used car market, it is expected that buyers will give more consideration to the future value of their cars. And petrol and diesel will likely have a much lower value after 2030.

UK Used Car Sales

The details in the data reflect the information given to potential buyers in the website adverts.

Buyers typically want to know the road tax of a used car, which varies based on a combination of year registered, emissions and fuel type, with Electric cars currently exempt from road tax.

Overview

6,200 miles

839 miles above average ⓘ

⛽ Petrol	🚗 Hatchback	1.6L	Manual
3 doors	5 seats	1 owner ⓘ	✓ ULEZ ⓘ

Description

Black, 1 owner, £34,995 **Read more**

⚙ **Specification**	›

🧮 **Running costs**	›

Example of a used car advert from [AutoTrader](#)

Data Information

The sales team has pulled some data from the website listings from the last 6 months. All cars listed in this data were sold at the price listed.

I think there is also a lot of information missing that could be useful to us.

You can download the data from here.

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

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.

Column Name	Details
model	Character, the model of the car, 18 possible values
year	Numeric, year of registration from 1998 to 2020
price	Numeric, listed value of the car in GBP
transmission	Character, one of "Manual", "Automatic", "Semi-Auto" or "Other"
mileage	Numeric, listed mileage of the car at time of sale
fuelType	Character, one of "Petrol", "Hybrid", "Diesel" or "Other"
tax	Numeric, road tax in GBP. Calculated based on CO_2 emissions or a fixed price depending on the age of the car.
mpg	Numeric, miles per gallon as reported by manufacturer
engineSize	Numeric, listed engine size, one of 16 possible values

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 8-10 slides to present to the product manager. 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 product manager 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.

Sample Solution

You can find a sample solution from a published workspace and a recording.

The sample solution demonstrates one possible approach to the problem given.