CASTLERY

Data Science Take-Home Test

- Hello from Castlery! This test contains 3 questions.
- When in doubt, state your assumptions and continue.
- Showcase both your technical command and business acumen. Have fun!

Question 1 (10 marks): Service Fee Change

Use data from "Q1 Service Fee Change.csv". This dataset includes total orders, Room of Choice (ROC) orders and White Glove (WG) orders in selected cities from furniture retailer Z's United States home delivery operations for a period. On 2022-09-12, ROC/WG service fees for customers in Miami, FL and San Francisco, CA were changed from $50/$100 to $70/$150.

Your mission:

1. Is the service fee change successful? (Hint: define success criteria)
2. What else do you observe from the dataset? Use suitable tables and charts.
3. If you are asked to determine optimal ROC/WG service fees, how will you do it? What are your guestimates of the final numbers?

Please submit your answer in .ipynb or other suitable format.

Question 2 (10 marks): Demand Forecasting

Use data from "Q2 Demand Forecast.csv". This dataset includes Furniture retailer Z's historical monthly sales in a sample region, along with several key variables captured on the first day of the month,

- is_online_1st_day: if the sku is visible to ecommerce customers
- lead_time_1st_day: estimated waiting days till the product can be dispatched
- original_price_1st_day: listing price of the sku on the day
- sale_price_1st_day: discounted price of the sku on the day

Your mission:

1. Inspect and clean the dataset. Use relevant tables and charts.
2. Use a suitable time-series forecasting model to predict SKU level demand for the following 12 months, including expected accuracy and variability.
3. Use another model (deep learning if it was not used) to predict the same. Explain the pros and cons of the 2 approaches.
4. If you have access to the company's transactional data, how would you better prepare the input dataset? Do you foresee any limitations of your chosen model?

Please submit your answer in .ipynb or other suitable format.

Question 3 (10 marks): Los Angeles Deliveries

Furniture retailer Z works with a variety of logistics partners to manage home deliveries from local warehouses to customers' apartments/ houses. To further improve customer experience and support rapid growth, the company is considering launching a small delivery fleet in Los Angeles metro area* and will need your help to evaluate its feasibility.

* This refers to metro_name = "Los Angeles, CA". In United States, cities are more narrowly defined.

Use data from "LA Deliveries.csv" (cbm = cubic metre, population = number of residents living in the same zipcode area, lat/ lng = latitude and longitude of the address).

Other known facts: a typical delivery truck can fit at most 20 cbm of freight and works on Monday to Friday for 10 hours per day; average time to complete one delivery job (from parking to moving off to the next destination) is about 30 minutes; average driving speed is about 50 km/hr; last-mile hub's location = Anaheim, CA 92804, lat/lng: 33.81826/-117.97506.

1. What is the minimum number of trucks we need given the above information to ensure all deliveries can be completed strictly based on the original delivery dates?
2. Let's name the new trucks T1, T2, etc. Assign each delivery job to one of the new trucks and show why each truck can complete the assigned job every day.
3. Do you think Castlery should operate her own fleet now? Please state your reasons.
4. What other information could make you change your recommendation above?

Please submit your answer in .ipynb or other suitable format.