ReCell - Problem Statement

Submission type

:

File Upload

Due Date

:

Jun 11, 3:30 AM

Total Score

:

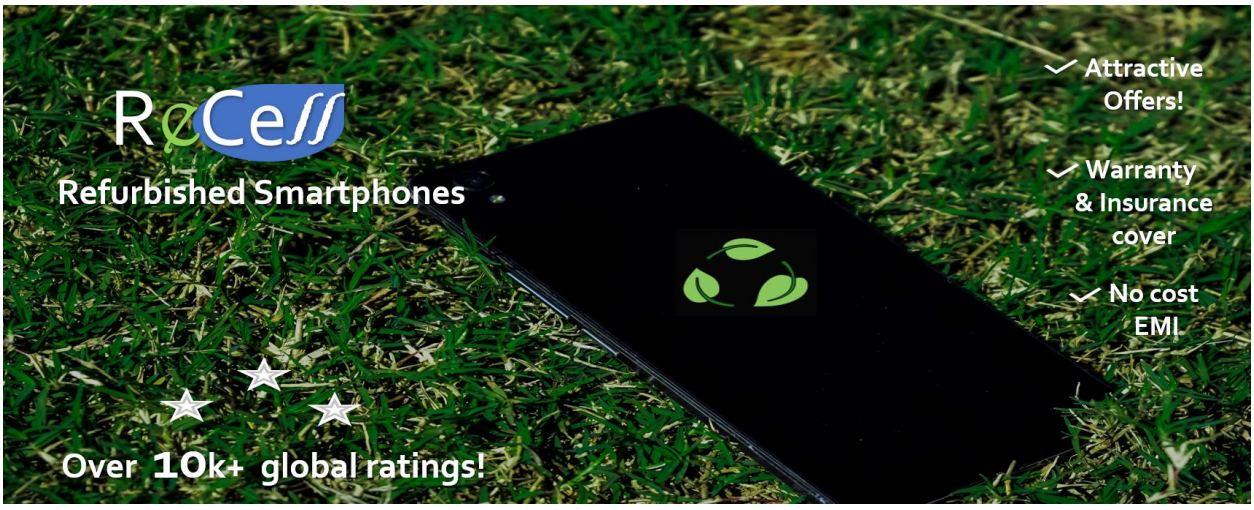
60

Available from

:

May 19, 5:30 PM

Description



Context

Buying and selling used phones and tablets used to be something that happened on a handful of online marketplace sites. But the used and refurbished device market has grown considerably over the past decade, and a new IDC (International Data Corporation) forecast predicts that the used phone market would be worth $52.7bn by 2023 with a compound annual growth rate (CAGR) of

13.6% from 2018 to 2023. This growth can be attributed to an uptick in demand for used phones and tablets that offer considerable savings compared with new models.

Refurbished and used devices continue to provide cost-effective alternatives to both consumers and businesses that are looking to save money when purchasing one. There are plenty of other benefits associated with the used device market. Used and refurbished devices can be sold with warranties and can also be insured with proof of purchase. Third-party vendors/platforms, such as Verizon, Amazon, etc., provide attractive offers to customers for refurbished devices. Maximizing the longevity of devices through second-hand trade also reduces their environmental impact and helps in recycling and reducing waste. The impact of the COVID-19 outbreak may further boost this segment as consumers cut back on discretionary spending and buy phones and tablets only for immediate needs.

Objective

The rising potential of this comparatively under-the-radar market fuels the need for an ML-based solution to develop a dynamic pricing strategy for used and refurbished devices. ReCell, a startup aiming to tap the potential in this market, has hired you as a data scientist. They want you to analyze the data provided and build a linear regression model to predict the price of a used phone/tablet and identify factors that significantly influence it.

Data Description

The data contains the different attributes of used/refurbished phones and tablets. The data was collected in the year 2021. The detailed data dictionary is given below.

Data Dictionary

- brand_name: Name of manufacturing brand
- os: OS on which the device runs
- screen_size: Size of the screen in cm
- 4g: Whether 4G is available or not
- 5g: Whether 5G is available or not
- main_camera_mp: Resolution of the rear camera in megapixels
- selfie_camera_mp: Resolution of the front camera in megapixels
- int_memory: Amount of internal memory (ROM) in GB
- ram: Amount of RAM in GB
- battery: Energy capacity of the device battery in mAh
- weight: Weight of the device in grams
- release_year: Year when the device model was released
- days_used: Number of days the used/refurbished device has been used
- normalized_new_price: Normalized price of a new device of the same model in euros
- normalized_used_price: Normalized price of the used/refurbished device in euros

Submission Guidelines

1. There are two ways to work on this project:

i. Full-code way: The full code way is to write the solution code from scratch and only submit a final Jupyter notebook with all the insights and observations.

ii. Low-code way. The low-code way is to use an existing solution notebook template to build the solution and then submit a business presentation with insights and recommendations.

The primary purpose of providing these two options is to allow learners to opt for the approach that aligns with their individual learning aspirations and outcomes. The below table elaborates on these two options.

Submission type	Who should choose	What is the same across the two	What is different across the two	Final submission file [IMP]	Submission Format
Full-code	Learners who aspire to be in hands-on coding roles in the future focussed on building solution codes from scratch	Perform exploratory data analysis to identify insights and recommendations for the problem	Focus on code writing: 10-20% grading on the quality of the final code submitted	Solution notebook from the full-code template submitted in .html format	.html
Low-code	Learners who aspire to be in managerial roles in the future-focussed on solution review, interpretation, recommendations, and communicating with business		Focus on business presentation: 10-20% grading on the quality of the final business presentation submitted	Business presentation in .pdf format with problem definition, insights, and recommendations	.pdf

Please follow the below steps to complete the assessment. Kindly note that if you submit a presentation, ONLY the presentation will be evaluated. Please make sure that all the sections mentioned in the rubric have been covered in your submission.

i. Full-code version

- Download the full-code version of the learner notebook.
- Follow the instructions provided in the notebook to complete the project.

- Clearly write down insights and recommendations for the business problems in the comments.
- Submit only the solution notebook prepared from the learner notebook [format: .html]

ii. Low-code version

- Download the low-code version of the learner notebook.
- Follow the instructions provided in the notebook to complete the project.
- Prepare a business presentation with insights and recommendations to the business problem.
- Submit only the presentation [format: .pdf]

2. Any assignment found copied/plagiarized with other submissions will not be graded and awarded zero marks.

3. Please ensure timely submission as any submission post-deadline will not be accepted for evaluation.

4. Submission will not be evaluated if

- it is submitted post-deadline, or,
- more than 1 file is submitted.

Best Practices for Full-code submissions

- The final notebook should be well-documented, with inline comments explaining the functionality of code and markdown cells containing comments on the observations and insights.
- The notebook should be run from start to finish in a sequential manner before submission.
- It is important to remove all warnings and errors before submission.
- The notebook should be submitted as an HTML file (.html) and NOT as a notebook file (.ipynb).
- Please refer to the FAQ page for common project-related queries.

Best Practices for Low-code submissions

- The presentation should be made keeping in mind that the audience will be the Data Science lead of a company.
- The key points in the presentation should be the following:
    - Business Overview of the problem and solution approach
    - Key findings and insights which can drive business decisions
    - Business recommendations
    - Focus on explaining the key takeaways in an easy-to-understand manner.
    - The inclusion of the potential benefits of implementing the solution will give you the edge.
- Copying and pasting from the notebook is not a good idea, and it is better to avoid showing codes unless they are the focal point of your presentation.
- The presentation should be submitted as a PDF file (.pdf) and NOT as a .pptx file.
- Please refer to the FAQ page for common project-related queries.

Happy Learning!

guide (Rubric) - ReCell Project Rubric

Criteria	Points
Exploratory Data Analysis - Problem definition, questions to be answered - Data background and contents - Univariate analysis - Bivariate analysis - Insights based on EDA	12
Data preprocessing - Duplicate value check - Missing value check and treatment - Outlier check (treatment if needed) - Feature engineering (if needed) - Data preparation for modeling	8
Model building - Linear Regression - Build the model and comment on the model statistics - Display model coefficients with column names	10
Testing the assumptions of linear regression model - Perform tests for the assumptions of the linear regression - Comment on the findings from the tests	10
Model performance evaluation Evaluate the model on different performance metrics	6
Actionable Insights & Recommendations - Comments on significance of predictors - Key takeaways for the business	6
Presentation/Notebook - Overall quality - Structure and flow - Crispness - Visual appeal - Conclusion and Business Recommendations OR - Structure and flow - Well commented code - Conclusion and Business Recommendations	8
Points	60