[ Sahithi Priya Rathod] [ Kandanelly ]

[3.416] [ California State University, Eastbay ] Questions with Time Spent:

Question 1 [ 14mins ]

Question 2 [ 12mins ]

Question 3 [ 15mins ]

Question 4 [ 20mins ]

Question 5 [ 2 hours ]

Question 6 [ 38 mins ]

Question 7 [ 42 mins ]

Question 8 [ 30 mins ]

Question 9 [ 01 hours ]

Question 10 [ 24 mins ]

Question 11 [ 36 mins ]

**Question 1**

Total Units Utilized per Month Time spent: 14mins

|  |  |  |  |
| --- | --- | --- | --- |
| **Sum of Units** | Column Labels |  | |
| Row Labels | Med A | Med B | Grand Total |
| 2012-07 | 4303700 |  | 4303700 |
| 2012-08 | 4477100 |  | 4477100 |
| 2012-09 | 849900 | 535 | 850435 |
| 2012-10 |  | 393 | 393 |
| 2012-11 | 75300 | 420 | 75720 |

2012-12 10200 1 10201

### Grand Total 9716200 1349 9717549

**Question 2**

No of patients receiving Med A and Med B from July to November.

A green and white chart

Description automatically generated

**Question 3**

Average total monthly dose per patient for each medication.

Time Spent: 15mins

Average of MED A in July = 5129.559 and MED B = N/A Average of MED A in August = 5645.77554 and MED B Average of MED A in September = 5311.875 and MED Average of MED A in October = N/A and MED B = 5.78 Average of MED A in November = 10757.1429

and MED B=5.6

**Question 4** Time Spent: 20mins

No of patients switched from Med A to Med B

Row Labels Sum of Switched

Sep 67

Oct 0

Nov 0

### Grand Total 67

First-Time Med B Users

M

|  |  |
| --- | --- |
| Row Labels | Sum of First time |
| Sep | 5 |
| Oct | 9 |
| Nov | 12 |

### Grand Total 26

**Question 5** Time Spent: 2hrs

Average number of weeks

A green rectangle with white text

Description automatically generated

**Approach 1:** Identified patients who transitioned from Med A to Med B. Subsequently, calculated the total average number of weeks each patient had been using Med A from the start of their treatment.

**Approach 2:** Identified the number of patients who switched from Med A to Med B and calculated the average number of weeks each patient used Med A during the period from September to November.

**Question 6** Time Spent: 38mins

Refer Question 6 A&B Sheet

**Question 7** Time Spent: 42mins

7202.52$ is the break even price point for med B

**Question 8** Time Spent: 30mins

Refer Question8a,b sheet

**Question 9** Time Spent: 01 hours

Refer Question 9 sheet

**Question 10** Time Spent: 24mins

Refer Question 10 sheet

**Question11** TimeSpent:36mins

Key Insights

- Medication A:

- Total LAB B Value: 181,477.1

- Higher total LAB B values indicate that Medication A is strongly associated with increased LAB B values.

- Medication B:

- Total LAB B Value: 22,726.5

- Lower total LAB B values compared to Medication A.

Impact on Breakeven Price Point

The breakeven price point is the price at which the costs associated with medications are equal to the revenue generated from their use. Given that LAB B values are generally associated with more usage of medications:

1. Higher LAB B Values with Medication A:

- If more of Medication A is associated with higher LAB B values, it suggests that Medication A has a stronger impact or effectiveness on LAB B.

- This means that the cost of Medication A might need to be justified by its higher effectiveness in producing desirable LAB B outcomes.

- The breakeven price point for Medication A could be higher due to its higher contribution to LAB B values.

2. Lower LAB B Values with Medication B:

- Medication B shows a lower total LAB B value, implying it might be less effective or used in lesser quantities compared to Medication A.

- The breakeven price point for Medication B could be lower, considering its lower impact on LAB B values.

Conclusion

To determine the exact breakeven price points, we would need detailed cost and revenue data associated with both medications. However, based on the given data:

- Medication A would likely have a higher breakeven price point due to its significant association with higher LAB B values.

- Medication B would have a lower breakeven price point due to its lesser association with LAB B values.

**SAHITHI PRIYA RATHOD KANDANELLY**

+1 510 329 8683 | [sahithipriya2109@gmail.com](mailto:sahithipriya2109@gmail.com) |Hayward, 94542, California | [linkedin.com/in/sahithi-priya](http://www.linkedin.com/in/sahithi-priya)

https://sahithis-portfolio.jimdosite.com

**EDUCATION**

**California State University, Eastbay Hayward, California, USA**

Master of Science in Business Analytics (STEM) Aug 2022 to May 2024

**Chaitanya Bharathi Institute of Technology, Gandipet, India Hyderabad, Telangana, INDIA**

Bachelor of Engineering in Electrical Engineering Jul 2018 to June 2022

**SKILLS**

**Programming Languages:** Python, SQL, R, C, Matlab.

**Data Analysis & Visualization:** Tableau , PowerBI , Matplotlib , seaborn , D3.js

**Data Engineering & Big Data:** Talend, AWS, Google BigQuery ,Amazon Glue Studio, Redshift, EMR, Hive, Hadoop, Apache Spark, PySpark **Machine Learning & Statistical Analysis:** Pandas, PyTorch, TensorFlow, NumPy, Scikit-learn, Regression Analysis, ANOVA

**EXPERIENCE**

**Aggress Tech Solutions Private Limited**  **Hyderabad, Telangana, INDIA**

**Data Analyst Intern Dec 2021-May2022**

**(Drive Link:** [https://drive.google.com](https://drive.google.com/file/d/13jjMUI065_HfDXT_GEC0tnzPG_QTTxNo/view?usp=sharing))

* Analyzed financial data and loan applications to assess creditworthiness and recommend loan approvals, resulting in a 20% increase in data accuracy.
* Engineered & executed innovative data cleansing protocols utilizing Python & SQL, streamlining loan processing duration by 15%.
* Created data visualizations and reports to communicate insights to stakeholders, optimizing decision-making processes. Utilized SQL queries and data analysis tools to extract and manipulate data.
* Supported the development of machine learning models for loan risk assessment, contributing to the identification of new trends in loan applications that led to a 10% improvement in risk assessment accuracy.

**PROJECTS**

**Sales Forecasting** <https://github.com/sahithipriya21/sales-forecasting>

* Participated in Kaggle challenge predicting quarterly sales for 75 customers of a steel manufacturer's base.
* Leveraged one-hot encoding to handle categorical variables and integrated economic indicators for enhanced predictive modeling.
* Secured a second-place submission score(MAE) of 643.84 getting by implementation of linear regression and random forest algorithms including Neural Networks that improved sales forecasting techniques and decision-making processes.

**SF Real Estate Analysis** <https://github.com/sahithipriya21/redfin>

* Transformed raw data from Redfin (2,500+ rows) using Python (ScrapeFly, NumPy, pandas, seaborn, matplotlib, scikit-learn)
* Cleaned non-numeric values, explored correlations, and implemented KNN classifier achieving 63.30% accuracy in categorizing bedrooms based on price and size.
* Verified positive correlation between property size and price, informing strategic purchasing decisions for homebuyers.

**Monthly Wholesale Trade Forecasting for Motor Vehicle and Supplies.** <https://github.com/sahithipriya21/timeseries>

* Applied regression-based models, Holt-Winter's Exponential Smoothing Model, and ARIMA models in R to analyze 30 years of data, unveiling seasonal patterns and trends, providing data-driven insights for strategic decision-making.
* Developed forecasting models, achieving RMSE of 720.943 and MAPE of 2.114.
* Confirmed Holt-Winters model with Automatic Selection as optimal, yielding lowest RMSE and MAPE values.

**Big Data Analytics and Processing Project with NCDC Records** <https://github.com/sahithipriya21/bigdata>

* Developed Python, PySpark, and MRJob applications to analyze large volumes of \*.gz files containing NCDC records. Tasks included wind direction averaging, sky ceiling height range calculation, and visibility distance extraction for USAF weather stations.
* Utilized Hadoop ecosystem tools such as Hadoop streaming, Pig, and Hive for scalable data processing and analysis. Leveraged Amazon S3 for storing project data and Amazon EMR for processing MapReduce jobs, optimizing efficiency and scalability with AWS cloud services.

**Smarter Handoffs: Machine Learning Optimizes Network Switching** <https://github.com/sahithipriya21/Handoff>

* Analyzed a dataset containing network parameters (signal strength, bandwidth, etc.) and handoff labels.
* Implemented three machine learning models: Logit model, Classification Tree, and Random Forests. Trained and evaluated the models using a 60/40 training-validation split. Achieved an accuracy of 88.75% with the Random Forest model, outperforming the other models. Identified key factors influencing handoff decisions, including signal strength, bandwidth, and signal-to-noise ratio.

**CERTIFICATIONS**  [https://drive.google.com/drive/u/2/folders/1eBmyyjUiYBHoOLTKxxI1pCu-xxsz\_b8n](https://drive.google.com/DRIVE/U/2/FOLDERS/1EBMYYJUIYBHOOLTKXXI1PCU-XXSZ_B8N)

[IBM DATA ANALYST](https://drive.google.com/drive/folders/1eBmyyjUiYBHoOLTKxxI1pCu-xxsz_b8n?usp=drive_link)[DATA ANALYSIS WITH PYTHON](https://drive.google.com/drive/folders/1eBmyyjUiYBHoOLTKxxI1pCu-xxsz_b8n?usp=drive_link)