

agathos

Data Analyst Project



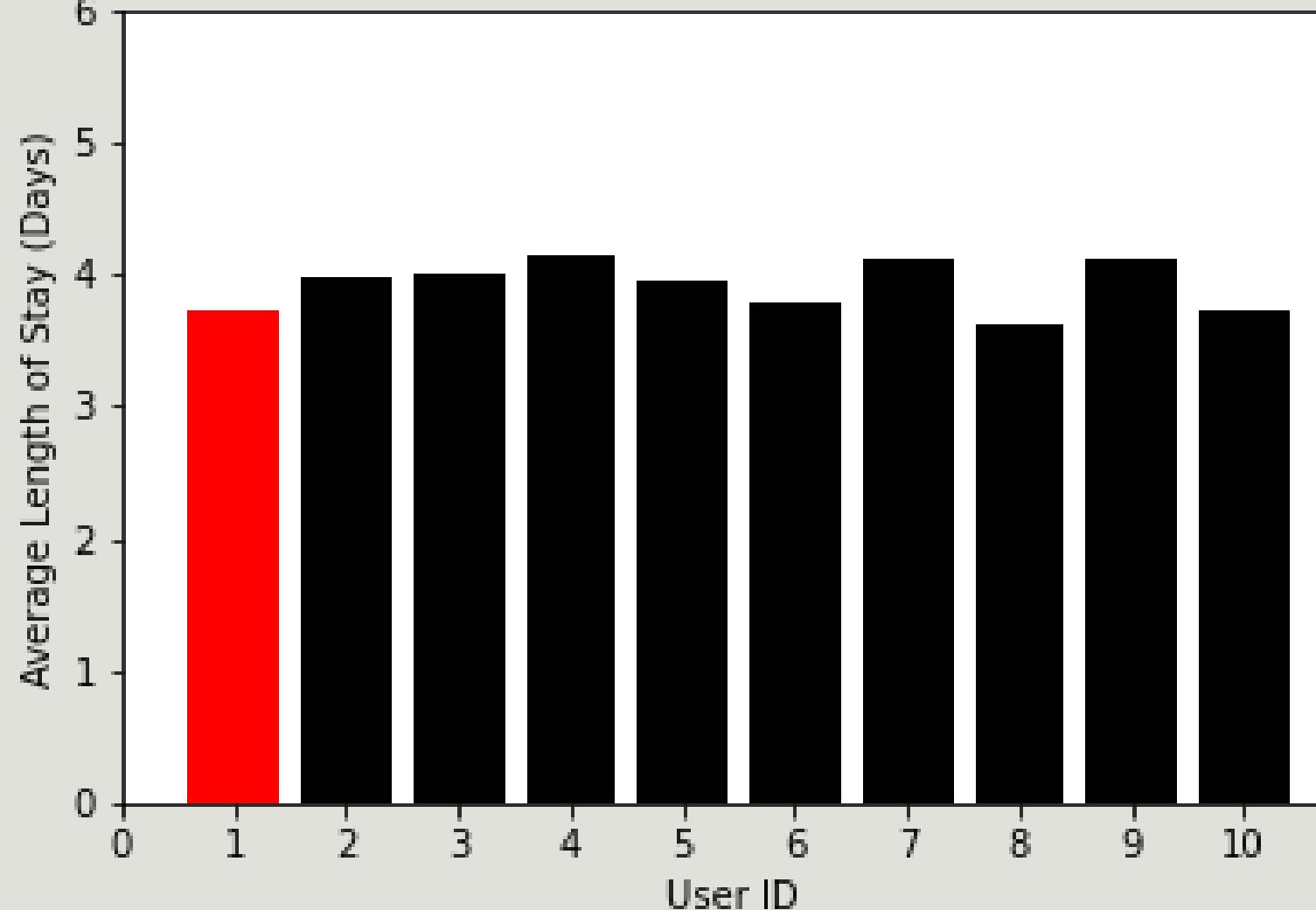
Problem Statement

A user (`user_id=1`) has complained that their monthly average length of stay (LOS) is higher than they expected when compared to their peers. In previous months, they had a much lower average LOS compared to their peers.

Given the group's set of hospital visits and their LOS for the current and previous month:

- Identify the root cause for this user's complaint.
- Determine and implement a fairer way to measure user average LOS.

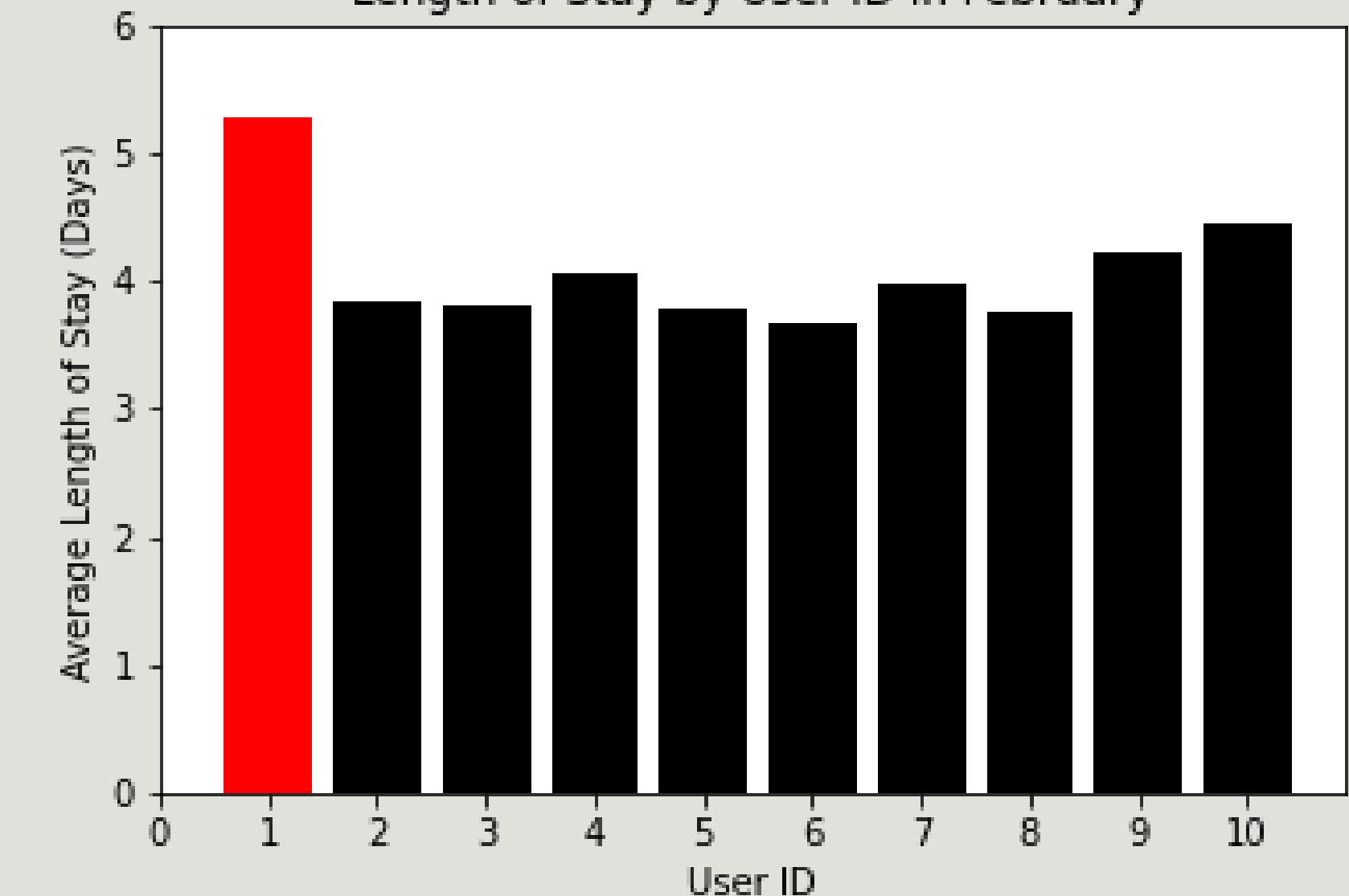
Length of Stay by User ID in January



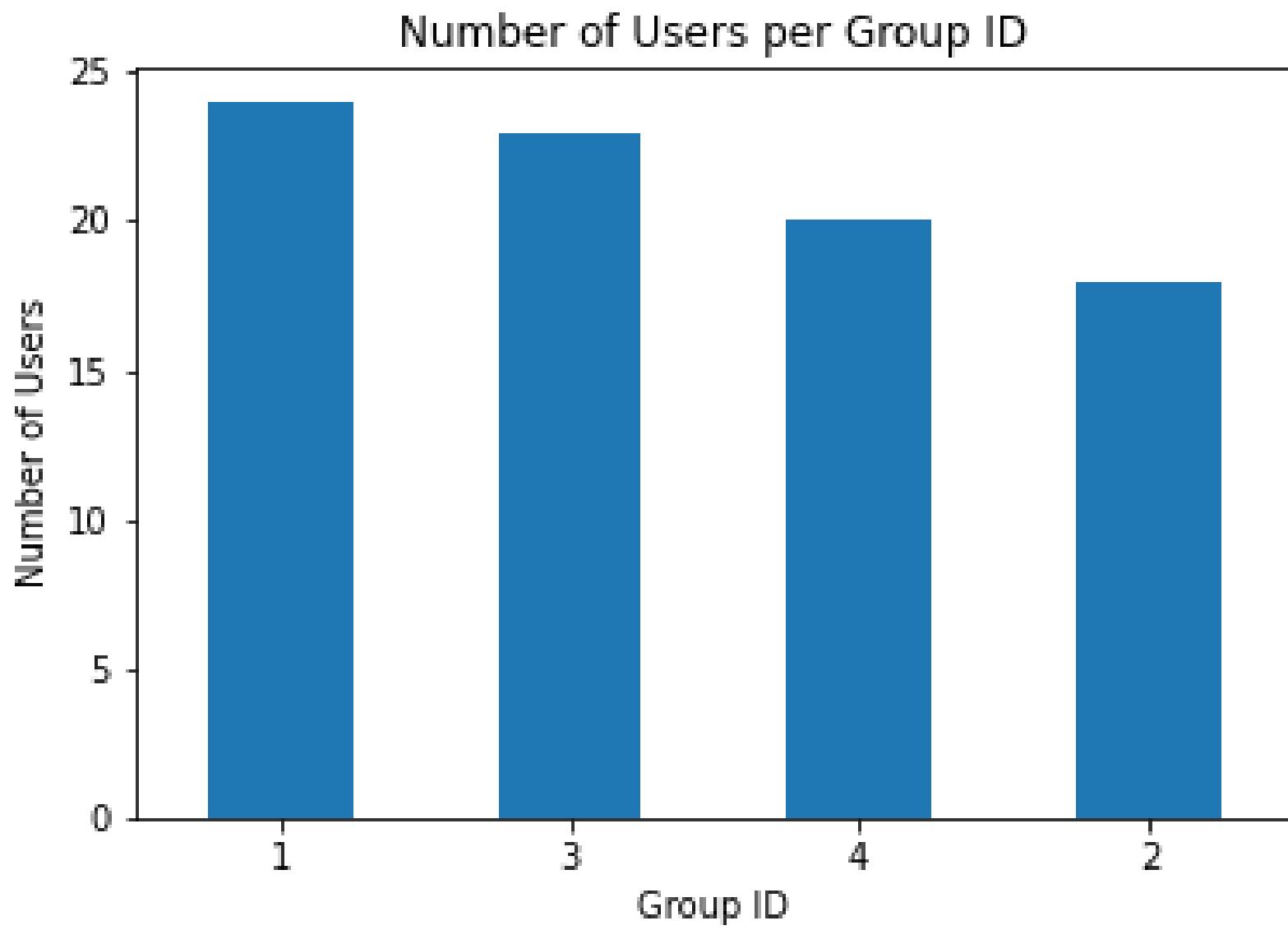
User_ID 1's average length of stay in January is consistent with other User_IDs in this month.

User_ID 1 is displayed in red, clearly illustrating a higher length of stay in February compared to other users.

Length of Stay by User ID in February



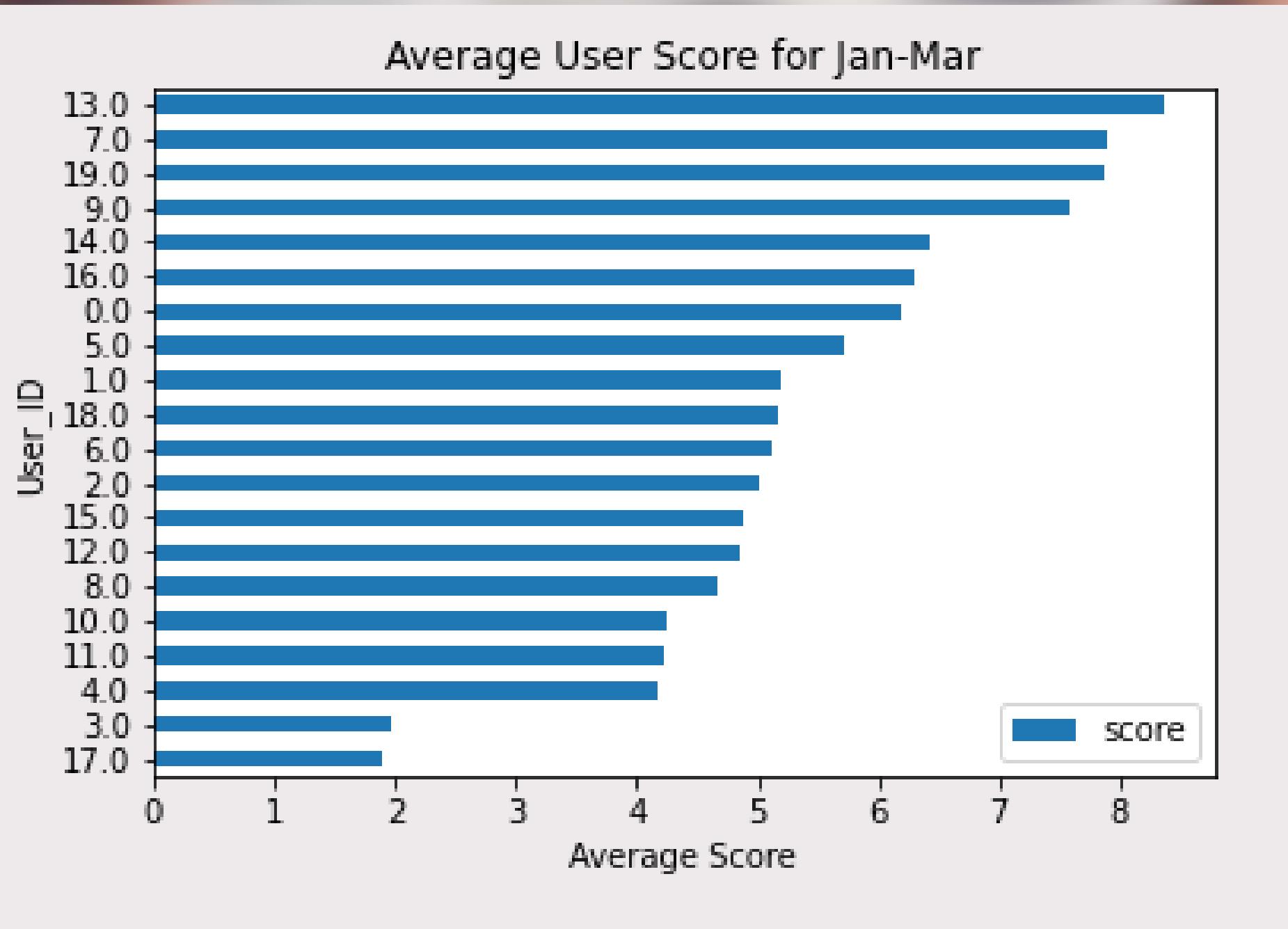
User_ID 1 is assigned to Group_ID 1, which has the most users (24) across the four groups.



- Given Group 1 (which contains User_ID 1) has the most users, it is likely this is the root cause for a higher average length of stay when compared to other groups with less user_ids.
- To implement a fairer way to measure user average length of stay, it is important to look at how users are grouped together and that sample sizes are comparable in size to ensure a larger group size is not inherently driving the average length of stay to increase.

Problem Statement

Given the following 3 months of monthly user scores, identify users that we could confidently say are bottom performers for this 3 month period and describe your reasoning.



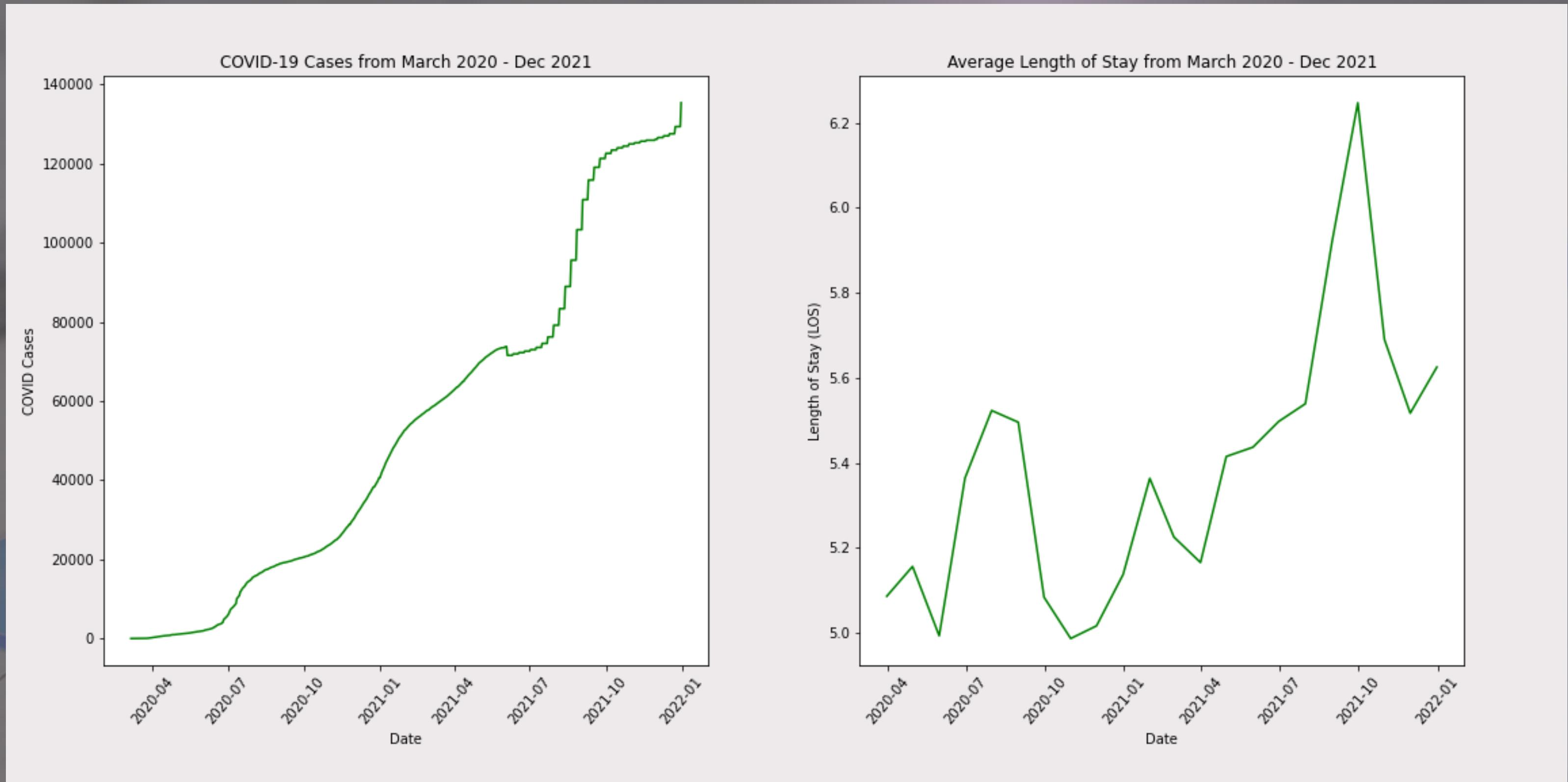
When visualizing the average user scores, it is clear to see which users are on the top, and which are bottom performers.

The following User_IDs are the bottom 3 performers: 17, 3, and 4.

User_IDs 17 and 3 have a notably lower performance rate as compared to the group as a whole. While User_ID 4 is in the bottom three performers (with an average score of 4.18,) their performance is not as low as User IDs 17 and 3 (1.88 and 1.96, respectively.)

Problem Statement

We observe that our client's average length of stay (LOS) for all hospital visits has had much more variation since the outbreak of COVID-19. Given a monthly measure of average LOS, at the client and cumulative COVID census for the same period, determine whether COVID is a plausible contributing factor.



When comparing COVID cases and average length of stay (LOS) over the same time period, it is plausible that COVID-19 contributed to the variability and rise in LOS. While the LOS varies greatly and COVID cases steadily increase, it is important to clarify that COVID could be an impacting factor though not the sole cause of the variance in LOS.



GET IN TOUCH

Lindsey Morales

LinkedIn: <https://www.linkedin.com/in/lindsey-morales->

GitHub: <https://github.com/valputla>



724-255-3990



lindsey-morales.squarespace.com/



lindseyvmorales@gmail.com



*Scan the QR code for quick
access to the github repository
associated with this
presentation.*