Interview - Assignment

> 💡 Please note that we expect you to share your work as python notebooks. You can either choose to work in Colab, Streamlit, Dash or somewhere where we can look at your work easily.
>
> Please keep your answers short and to the point. Please write clear code and keep things simple.

Question 1

Dataset: You are provided a CSV file in the below format.

```
pgdate_part, merchantid, count, sum
4, 10, 7871, 812381.01
```

Please write a script to visualize and present the new, churn, expansion and contraction GMV for these merchants.

Identify and build cohorts from this data and plot monthly retention for these cohorts?
Tip: think about how would you build cohorts for a B2B payments business.
You can read more about SaaS metrics here.

Question 2

Dataset: You are provided a CSV file in the below format.

```
upi_id, merchantid, txnId, txtime, txStatus, amount, remitterBank
12312, 875, 78612312,2022-10-11 09:08:34, SUCCESS, 100.0, SBI
12315, 990, 78612414,2022-10-11 09:10:33, FAILED, 100.0, ICICI Bank
12319, 190, 78613411,2022-10-11 09:11:12, USER_DROPPED, 10.0, NA
12325, 875, 78612412,2022-10-11 09:11:13, INCOMPLETE, 100.0, NA
```

Please write a script to identify if any `remitterBank` is facing a downtime which is leading to higher payment failure for that bank? Points to think -

1. How would you define Success rate?

2. How do you build a baseline success rate? What parameters are used?

3. How do you build a model to identify which remitter bank is facing a downtime?