



Question:

- Budget available for project?
- What is the typical size of each event in Dataset A?
 - 10000event per sec, 1million row of data in one event, google cloud are able to handle this size of data.
- Do we need to handle late arriving data?
 - I used window method in my design.
- Dataset B is store in google cloud storage or Big Query?
 - I assume is store in google cloud storage for my design.
- What strategy should we use for retrying fail event?
 - I used watermark as the solution for my design.
- How long should data be stored? if the data needed for analysis is past 32 days then data more than 32 days can be deleted or need to zip and store in cloud?
 - Either delete or zip the data to store in cloud to reduce the storage cost
- Who will need to access the dashboard?
 - This is use to set up IAM
- What is the action for Null?
 - I assume we keeping it for other analysis