Part 1: SQL Implementation

There is no sample dataset given for this assessment. Generate random dataset with your favorite scripting language.

Given a table user_events with columns:

- event_id (string)
- user_id (string)
- event_name (string) - possible values: 'PageView', 'Download', 'Install', 'Purchase'
- platform (string) - possible values: ios and android
- device_type (string)
- timestamp (timestamp)

Implement:

- Write an ANSI SQL query to create a funnel analysis showing the conversion rates between
- PageView → Download → Install.
- Show both the absolute numbers and conversion rates between steps.
- Only consider conversions that happen within 72 hours between steps
- Break down the funnel by platform.

Part 2: Data Modeling Questions

1. Looking at the events data above, how would you model this data in a production environment? Consider aspects like:

- Table structure
- Partitioning strategy
- What other tables might be needed?
- How would you handle data quality?

2. If we wanted to extend this analysis to include user attributes (like country, device type), what changes would you make to the data model?

3. What are potential issues with the current event tracking system that you can identify?

Part 3: Visualization

1. What tools would you recommend for visualizing this funnel data?
2. How would you design a dashboard to monitor these conversion rates over time?
3. What additional metrics or breakdowns would you include in the visualization?