

 **databricks**

<https://databricks.com> (Python)

  Import notebook

```
# Databricks notebook source
# -----
# Notebook: 03_gold_analytics
# Purpose : Build gold-level aggregated tables and views for analyti
#           curated silver data.
#
# Exam Coverage (Databricks Certified Data Engineer Associate – Exam
# - Section 3: Data Processing & Transformations
#   - Computing complex aggregations and metrics with Spark SQL / Da
#   - Designing gold tables as serving-ready datasets.
# - Section 5: Data Governance & Quality
#   - Ensuring aggregated revenue and trip metrics are consistent ar
#
# Key Practices
# - Aggregate by time grain (e.g. month) and business dimensions.
# - Use clear metric names and aliasing.
# - Assert non-negative revenue metrics where applicable.
# -----
```

Exam Focus – Complex Aggregations and Metrics (Section 3)

This notebook focuses on:

- Grouping silver data by time grain and dimensions (payment type, zones, companies).
- Computing metrics such as:
 - total_trips
 - total_revenue
 - avg_fare
- Writing aggregated results into gold Delta tables and exposing serving views.

Exam notes:

- Aggregations typically use `groupBy` + `agg` in DataFrames or `GROUP BY` in SQL.
- ;
- Use `count distinct` when a metric requires unique entities (e.g.