

CodeStates Project - CS1

E-Commerce Data Analysis

목차

- Olist 데이터셋
- DB, SQL
- Tableau 소개
- 대시보드
- 정리




DB, SQL

Olist 데이터셋

Olist 데이터셋 소개

- 브라질의 전자상거래 사이트 Olist에서 공개한 데이터셋
- 약 10만건의 주문 데이터와 상품, 결제, 판매자, 리뷰데이터 등 다양한 데이터가 존재



The screenshot shows the Kaggle page for the 'Brazilian E-Commerce Public Dataset by Olist'. The header includes the dataset title, a description of 100,000 orders, and the Olist logo. Below the header, there are tabs for 'Data', 'Code (192)', 'Discussion (43)', 'Activity', and 'Metadata'. A 'Download (126 MB)' button and a 'New Notebook' button are visible. The 'Usability' is 10.0, and the 'License' is CC BY-NC-SA 4.0. The 'Tags' section lists 'business', 'nlp', 'data visualization', 'exploratory data analysis', 'multiclass classification', and '1 more'. The 'Description' section provides a welcome message and details about the dataset's origin and features.

Brazilian E-Commerce Public Dataset by Olist
100,000 Orders with product, customer and reviews info

Olist and 2 collaborators • updated 3 months ago (Version 2)

[Data](#) [Code \(192\)](#) [Discussion \(43\)](#) [Activity](#) [Metadata](#) [Download \(126 MB\)](#) [New Notebook](#)

Usability 10.0 **License** CC BY-NC-SA 4.0 **Tags** business, nlp, data visualization, exploratory data analysis, multiclass classification and 1 more

Description

Brazilian E-Commerce Public Dataset by Olist

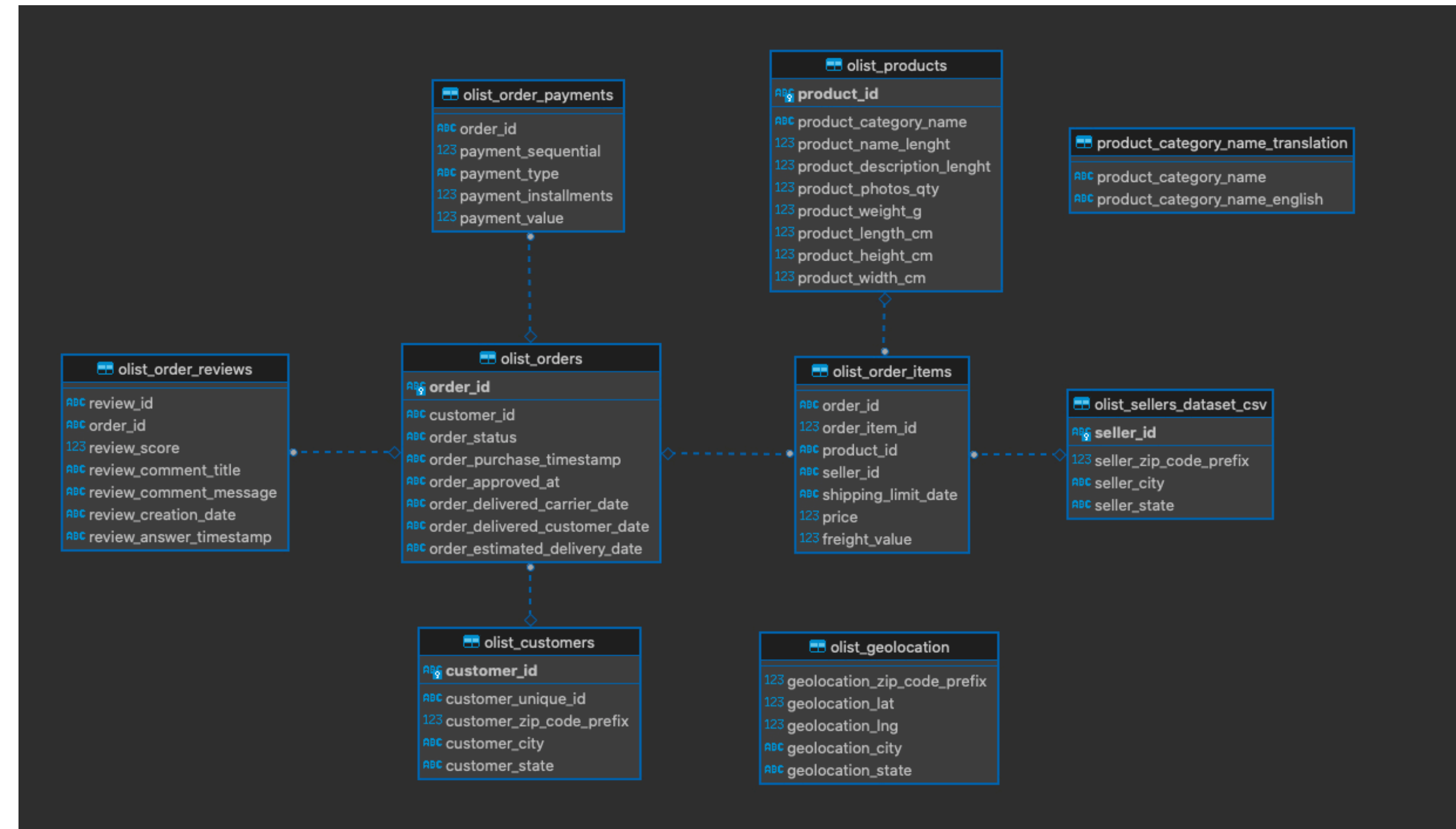
Welcome! This is a Brazilian ecommerce public dataset of orders made at [Olist Store](#). The dataset has information of 100k orders from 2016 to 2018 made at multiple marketplaces in Brazil. Its features allows viewing an order from multiple dimensions: from order status, price, payment and freight performance to customer location, product attributes and finally reviews written by customers. We also released a geolocation dataset that relates Brazilian zip codes to lat/lng coordinates.

This is real commercial data, it has been anonymised, and references to the companies and partners in the review text have been replaced with the names of Game of Thrones great houses.

DB, SQL

PostgreSQL

- 많이 사용되는 오픈소스 관계형데이터베이스인 PostgreSQL을 사용
- 데이터셋과 함께 주어진 엔티티관계도를 참고하여 스키마를 구성



DB, SQL

E-Commerce KPI

- 기간내 매출액(Revenue)
- 주문당 결제금액
- 객단가(ARPU)
- 판매자별 매출액
- 상품별 매출액



DB, SQL

PostgreSQL

- 쿼리문 실행

Tableau Dashboard

Tableau Dashboard

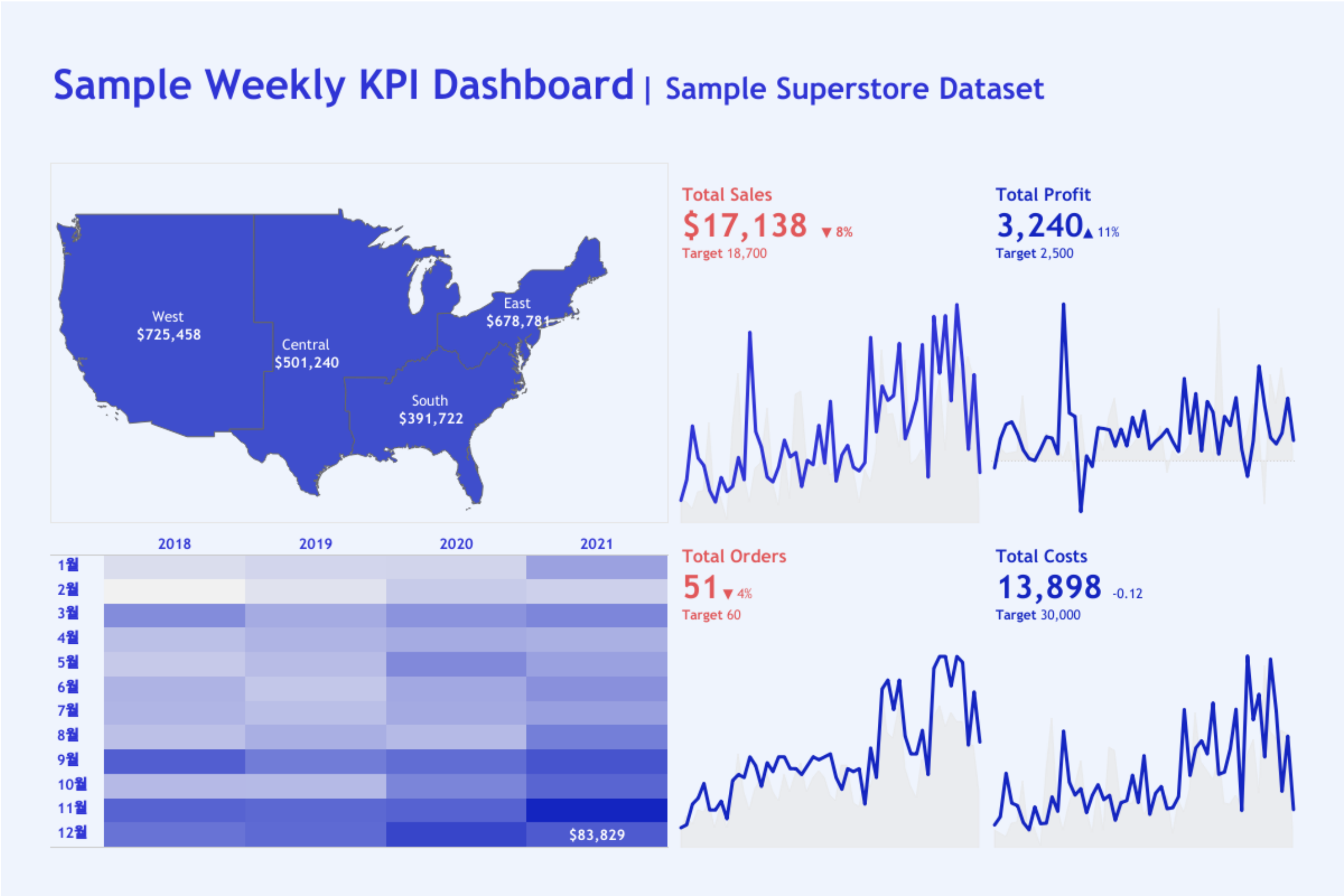
Tableau

- 많이 사용되고 기능이 풍부한 Tableau로 대시보드 제작
- Tableau 샘플 데이터로 데이터 시각화



Tableau Dashboard

Dashboard



정리

정리

느낀점

- SQL은 쿼리를 정확히 짜는 연습이 필요
- 대시보드는 정보를 효과적으로 전달하는게 중요