Problem Statement:

Assuming you are a Data/Business analyst at Target, you have been assigned the task of analyzing the given dataset to extract valuable insights and provide actionable recommendations.
What does 'good' look like?

1. Data types of all columns in the "customers" table
2. Time range (min/max) between which the orders were placed
3. Distinct cities & states among customers who actually placed orders (over the full period present)
4. Annual trend: number of orders per year (is there growth?)
5. Monthly seasonality: total orders by calendar month (across all years)
6. Time of day when orders are placed (Dawn/Morning/Afternoon/Night)
   -- Dawn: 0-6, Morning: 7-12, Afternoon: 13-18, Night: 19-23
7. Month-on-month number of orders per state (customer state)
8. Customer distribution across states (unique customers who ever ordered)
9. % increase in cost of orders from 2017 to 2018 (Jan–Aug only), using payments.payment_value
10. Total & average order price per state (price ≈ total payment per order)
11. Total & average order freight per state
12. Per-order delivery time and difference vs estimate (in days and intervals)
    -- time_to_deliver = delivered_customer_date - purchase_timestamp
    -- diff_estimated_delivery = delivered_customer_date - estimated_delivery_date
13. Top 5 states with highest & lowest average freight value
14. Top 5 states with highest & lowest average delivery time (days)
15. Top 5 states where delivery is fastest vs the estimate (positive = earlier than estimate)
16. Month-on-month number of orders by payment type
17. Number of orders by payment installments
18. Top 10 product categories by revenue (requires products.product_category_name)
19. Repeat purchase rate: customers with >=2 orders
20. Top 10 cities by number of orders, plus share of total orders