Background:

As a data analyst working for ABC Healthcare company, a claims dataset has been used that examines a population of patients enrolled in ABC's Accountable Care Organization care model, working with a High Deductible Health Plan (HDHP) insurance product to analyze claims paid on a population of patients enrolled in your ACO.

A dashboard has been created based on the specifications provided below in the Customer Service Request (CSR) form below.

This table contains the metadata needs to create a dashboard:

Field Name	Field Description
CLAIM_ID	Unique identifier for the insurance paid claims.
PATIENT_ID	De-identified ID for the patient.
PATIENT_STATE	The state in the USA where the patient currently resides.
PRODUCT_LINE	ACO includes Commercial (e.g. Blue Cross Blue Shield, Aetna, Cigna, etc.) and Medicare patients.
DATE_OF_SERVICE	The date of service on the insurance claim.
CPT_CODE	Current Procedural Terminology (CPT) code for the service rendered on the patient on the claim.
CPT_DESCRIPTION	Current Procedural Terminology (CPT) description for the service rendered on the patient on the claim.
PROVIDER_NPI	Providers National Provider Identification (NPI) number, uniquely assigned to every provider practicing medicine in the US.
PROVIDER_SPECIALTY	Provider's specialty indicator.
TOTAL_PAID_BY_INSURANCE	Provider bills for services typically split between the amount paid by insurance and the rest is collected as deductible from the patient. This field is the insurance paid amount.
PATIENT_DEDUCTIBLE	Provider bills for service is typically split between the amount charged to insurance and the rest is collected as deductible from the patient. This field is the patient paid deductible amount.

1) **First visualization:**

Create a horizontal bar graph showing the sum of total paid by insurance grouped by provider specialty and understand the key findings using various visualization techniques.

What are the two highest paid provider specialties?

The highest paid provider specialties are **Internal Medicine and Nurse Practitioner**.

Can you visually indicate these two directly on your visualization?

Yes, we can visually indicate the two highest paid provider specialties by distinguishing the horizontal bars of provider specialties using a color scheme and highlighting the top two bars as Blue and the rest as grey.

2) Second visualization:

Create a line graph showing the total paid by insurance by Year and Month, for Commercial vs. Medicare

Which product line appears to cost more for ABC Healthcare company? Medicare or Commercial?

Based on line graph, it seems like **Medicare** appears to cost more for the company.

Are there any months that look like outliers in terms of very high total paid by insurance?

The line graph representing the Medicare shows that there lies an outlier for the month of **February 2018** in which the total paid by the insurance is the highest.

Can you visually indicate any outlier months directly on your dashboard?

Yes, we can directly highlight the outlier on the dashboard using an Annotation mark.

3) Third visualization:

Create a scatter plot of total paid by insurance, vs. patient's deductible.

Remove outlier points by filtering only to observations where insurance paid is $2K or less

Does there appear to be a linear correlation between total paid by insurance and patient's deductible?

Yes, there exist a correlation between the total paid by insurance and Patient's deductible. There is a **weak-negative relation** between the two variables which is represented by a trend line. The relation is negative because as the amount paid by the patient increase, the total paid by insurance decreases.

4) Fourth Visualization

The 4th graph represents the Quarterly analysis of the amount spent by insurance in each state. Through this analysis, it was found that **Massachusetts (MA) has paid the highest in insurance** in the third quarter in the year 2017.

Can you visually indicate your finding or findings directly on your visualization?

A timeline graph is used to represent quarters in the year 2017 and 2018 and each state is represented using bubbles and a color scheme, where the size of the bubble shows the amount spent by the insurance. To highlight the state paid the highest insurance, an annotation mark is used to represent the State name, Quarter, Year and Total amount paid by the insurance.