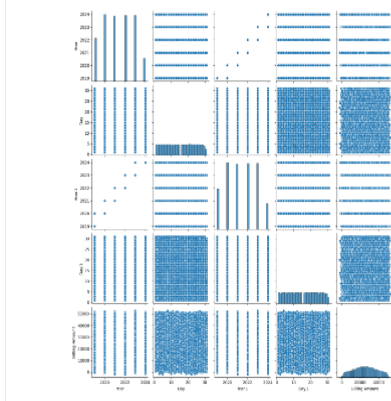| Name	Ameya Dabholkar
UID	2021300023
Batch	F

Dataset Link:

Dataset Description:

- **Name:** This column represents the name of the patient associated with the healthcare record.

- **Age:** The age of the patient at the time of admission, expressed in years.

- **Gender:** Indicates the gender of the patient, either "Male" or "Female."

- **Blood Type:** The patient's blood type, which can be one of the common blood types (e.g., "A+", "O-", etc.).

- **Medical Condition:** This column specifies the primary medical condition or diagnosis associated with the patient, such as "Diabetes," "Hypertension," "Asthma," and more.

- **Date of Admission:** The date on which the patient was admitted to the healthcare facility.

- **Doctor:** The name of the doctor responsible for the patient's care during their admission.

- **Hospital:** Identifies the healthcare facility or hospital where the patient was admitted.

- **Insurance Provider:** This column indicates the patient's insurance provider, which can be one of several options, including "Aetna," "Blue Cross," "Cigna," "UnitedHealthcare," and "Medicare."

- **Billing Amount:** The amount of money billed for the patient's healthcare services during their admission. This is expressed as a floating-point number.

- **Room Number:** The room number where the patient was accommodated during their admission.

- **Admission Type:** Specifies the type of admission, which can be "Emergency," "Elective," or "Urgent," reflecting the circumstances of the admission.

- **Discharge Date:** The date on which the patient was discharged from the healthcare facility, based on the admission date and a random number of days within a realistic range.

- **Medication:** Identifies a medication prescribed or administered to the patient during their admission. Examples include "Aspirin," "Ibuprofen," "Penicillin," "Paracetamol," and "Lipitor."

- **Test Results:** Describes the results of a medical test conducted during the patient's admission. Possible values include "Normal," "Abnormal," or "Inconclusive," indicating the outcome of the test.
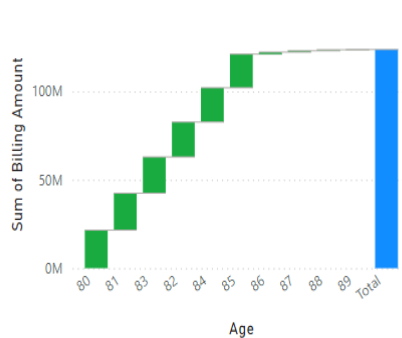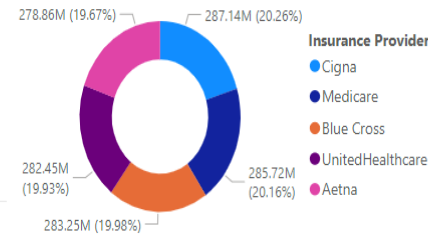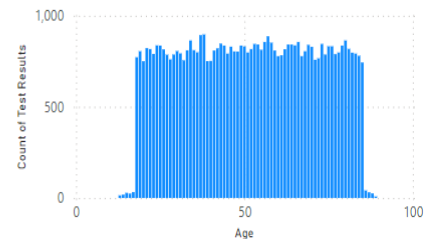
Report:

HealthCare Dataset EDA

This EDA explores the state of the healthcare industry in the USA. The analysis reveals that males and females are equally prone to diseases and hospital admissions. Obesity stands out as both the most prevalent and costliest condition to manage in the short term. A donut chart highlights Cigna as the dominant insurance provider in the sector. Additionally, the data shows that elderly individuals tend to have higher hospital bills compared to other age groups. The first half of the analysis presents the billing amounts across different age groups, categorized by day, month, quarter, and year.
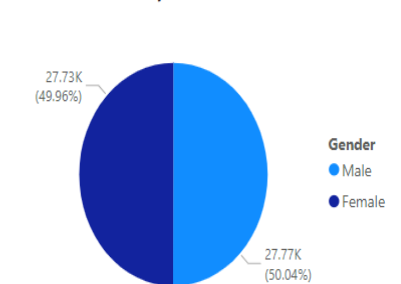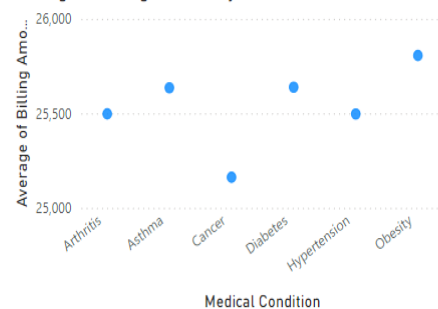
⬚ **How does the billing amount vary across different age groups, and which age group contributes the most to hospital billing?**

- **Answer**: The billing amount increases with age, and the elderly (around age 89) contribute the most to hospital bills. This indicates that older individuals tend to have higher healthcare costs.

⬚ **What is the distribution of test results across age groups, and are there any significant trends or patterns in the data?**

- **Answer**: The test results show a relatively even distribution across all age groups, with the highest counts being in middle-aged and older individuals. There doesn't appear to be any significant outliers, indicating consistent testing across the population.

⬚ **Which insurance provider has the largest share of hospital billing, and how does Cigna compare to other providers like Medicare and Blue Cross?**

- **Answer**: Cigna is the largest insurance provider, holding a slight lead over Medicare and other providers like Blue Cross and UnitedHealthcare. This indicates Cigna's dominance in covering medical expenses in the dataset.

⬚ **How does the average billing amount differ by medical condition, and why is obesity the most expensive condition to treat?**

- **Answer**: Obesity has the highest average billing amount compared to other conditions like arthritis, asthma, and diabetes. This could be due to the complexity and long-term management required for obesity-related treatments.

Is there an observable difference in the number of males and females admitted to hospitals, and does this reflect an equal susceptibility to diseases?

- **Answer**: The gender distribution is almost equal, with a slight majority of females (50.04%) compared to males (49.96%). This suggests that both genders are nearly equally susceptible to hospital admissions.

How do hospital bills differ by age on a quarterly, monthly, and yearly basis, and what trends can be inferred from the time-based analysis?

- **Answer**: The time-based analysis shows consistent increases in hospital bills as people age. Bills are higher in older age groups, which is consistent across different time intervals, indicating a trend of increasing healthcare costs with aging.

What is the relationship between the total billing amount and insurance providers, and how might this influence healthcare policy decisions?

- **Answer**: The total billing amount is distributed relatively evenly among the top five insurance providers, with Cigna leading. This could influence healthcare policy decisions, encouraging more competition among insurers to control costs, or focusing on preventative care to reduce high costs from chronic conditions like obesit