Summary of Insights

The dataset consists of two primary files Beneficiary Summary File and Inpatient Claims File.Given Datasets provides information on Medicare beneficiaries, including demographic details, chronic conditions, and financial data related to claims. By analyzing and merging these datasets we can derive valuable insights into the US healthcare system.

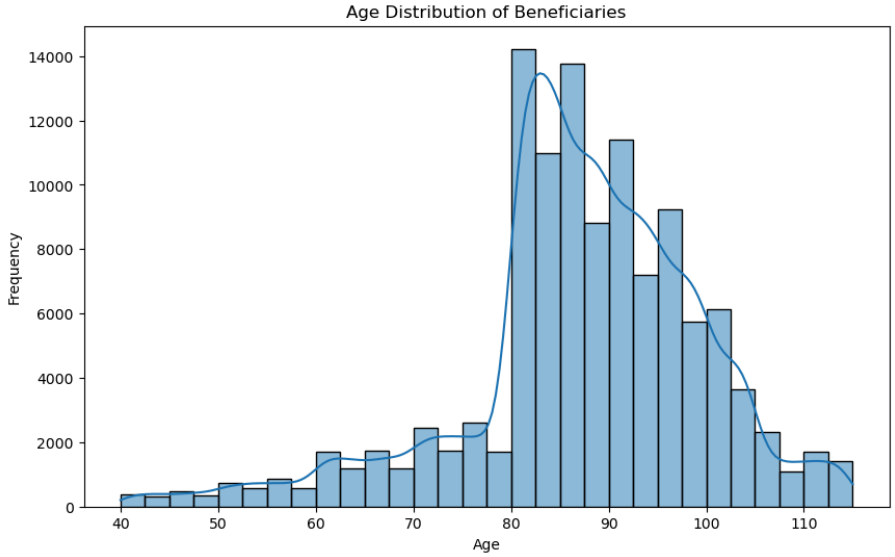
Beneficiary Summary File : ID, Date of Birth, Sex, Race, State Code, Disease Details.
Inpatient Claims File: ID, Claim_ID, Claim Dates(start, end), Claim Amount, Reimbursement Amount.

Steps and Insights in Analysis:

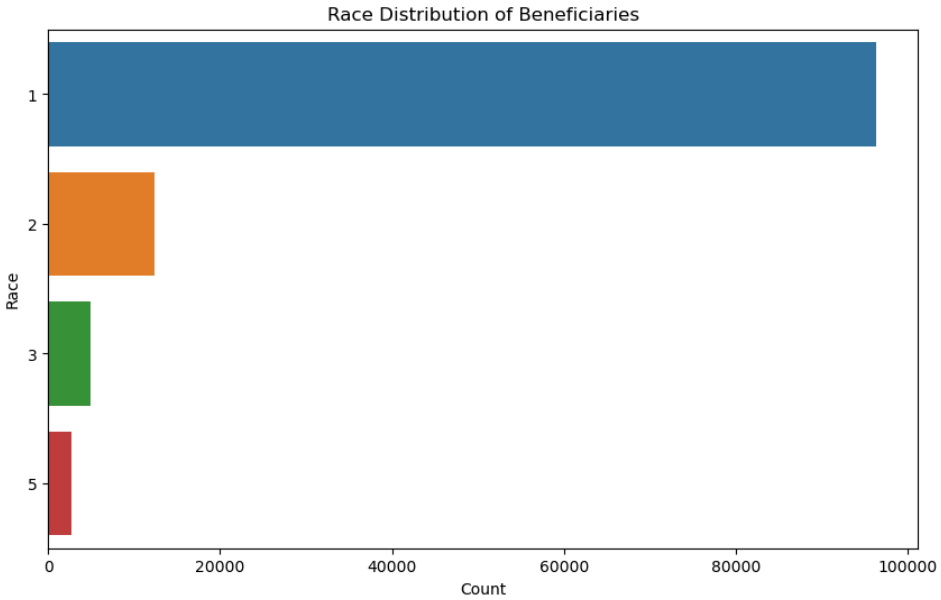
- Cleaning data by removing null values with mean
- Understanding Areas of analysis (As mentioned in description pdf)
- Demographic(Age, Race)
- Disease (Chronic Conditions)
- Financial (Claim, Reimbursement)
- Financial claim by race in Texas
- Visits and claim

Demographic Insights:

- Majority of beneficiaries are senior citizens, with significant numbers aged 75-85.



- Racial distribution shows that most beneficiaries are white.
- Higher proportion of white beneficiaries can be correlated to there general higher population. Eg: White: 60.1, Black: 12.2, Hispanic: 18.7, Other: 9.2 .

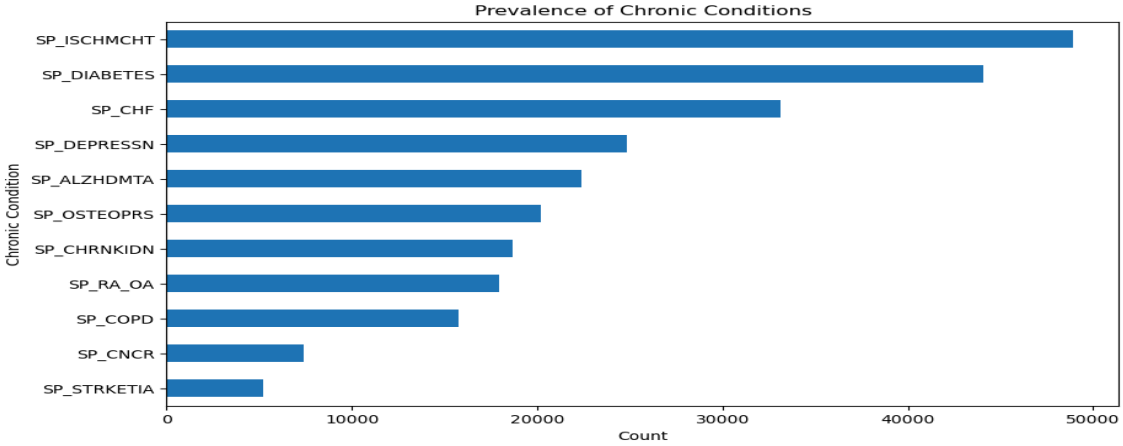


Categories

Code	Label	2008 Freq.	2008 Freq. in %	2009 Freq.	2009 Freq. in %	2010 Freq.	2010 Freq. in %
1	White	1,926,708	82.8	1,897,108	82.8	1,866,993	82.79
2	Black	247,723	10.65	244,068	10.65	240,294	10.66
3	Others	97,972	4.21	96,480	4.21	95,012	4.21
5	Hispanic	54,453	2.34	53,664	2.34	52,799	2.34
Total		**2,326,856**	**100**	**2,291,320**	**100**	**2,255,098**	**100**

Chronic Conditions Insights:

- Most common chronic conditions are Diabetes, Heart Failure, Ischemic Heart Disease, Alzheimer, Depression etc.
- Chronic conditions increases with age, particularly in the 75-84 age group.



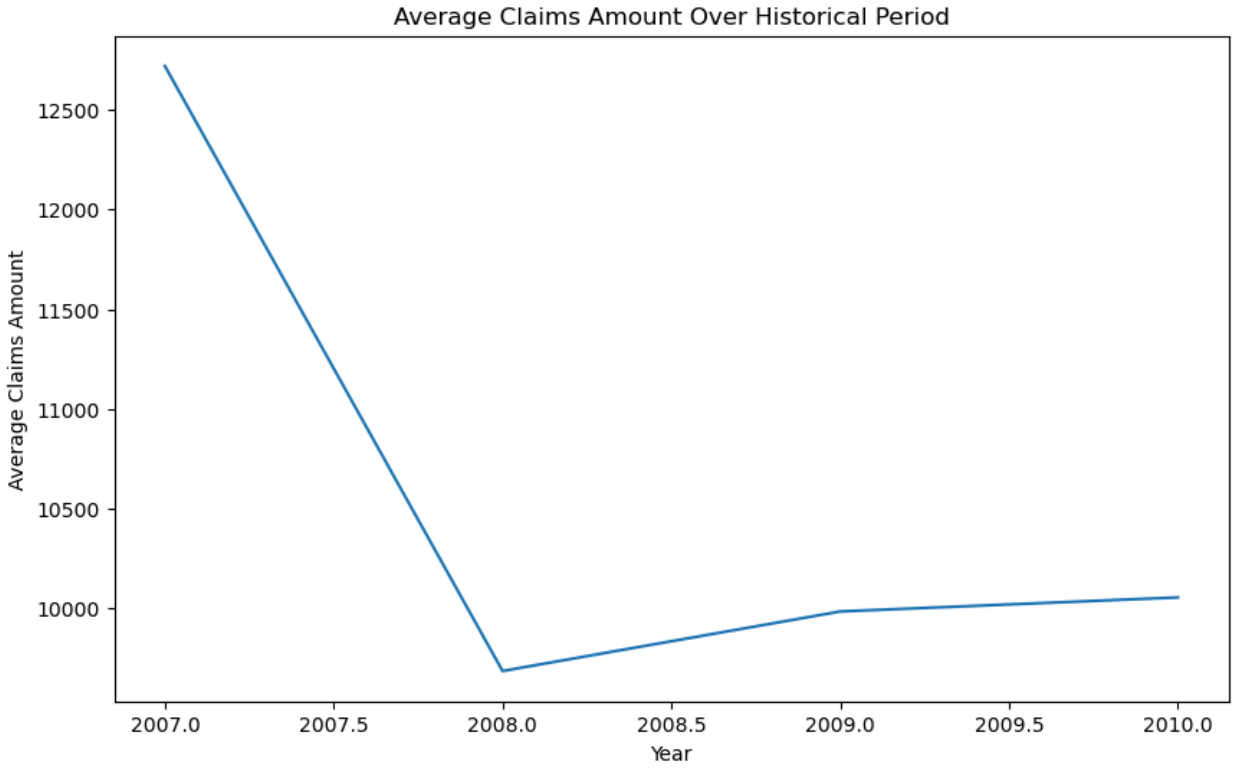
-

Financial Insights:

- The average claim amount is approximately $13,724.41, with a total of about $17.3 billion.
- The average reimbursement amount is approximately $11,550.09, with a total of about $14.6 billion.
- The reimbursement ratio varies by race, with potential disparities indicating inequities in healthcare costs and reimbursements.

	Race	Reimbursement_Ratio
3	White	3.111813
0	Black	3.172441
1	Hispanic	3.297672
2	Other	5.250800

- Sudden drop in Average Claims Amount Over Historical Period coincide with Great Recession of 2007 and 2008 which caused “Loss of Health Insurance” due to major Layoffs
- Reduced healthcare spending due to slowdown and people ignored necessary medical care due to concerns about affordability.



- Healthcare spending(claims) increase overtime indicate increase in utilization of facilities.

Hypothesis Testing:

- Null Hypothesis (H0H_0H0): There is no significant difference in the reimbursement ratios between different racial groups.
- Alternative Hypothesis (H1H_1H1): There is a significant difference in the reimbursement ratios between different racial groups.

Statistical Test:

- Use ANOVA (Analysis of Variance) to test for differences in the mean reimbursement ratios across multiple racial groups.
- If ANOVA shows significant differences, follow up with post-hoc tests to determine which groups differ from each other.

	sum_sq	df	F	PR(>F)
C(BENE_RACE_CD)	4.058851e+02	3.0	0.587337	0.623238
Residual	1.538045e+07	66769.0	NaN	NaN

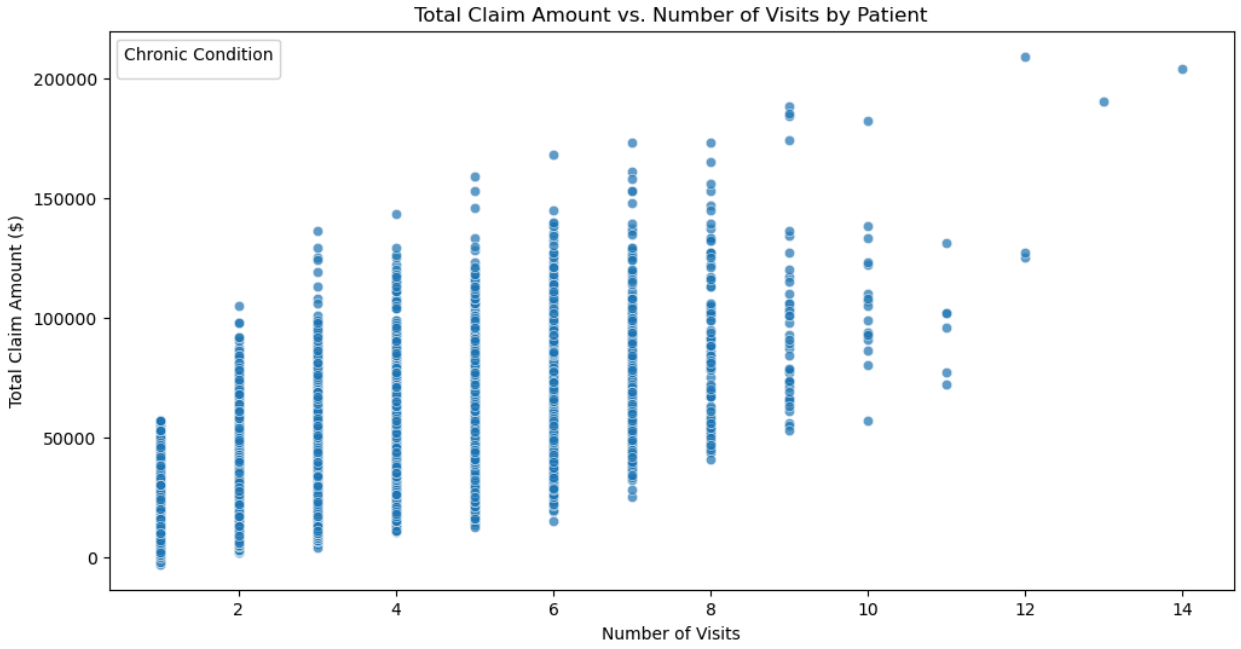
Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
1	2	-0.1686	0.8169	-0.6627	0.3255	False
1	3	-0.2182	0.9172	-1.0868	0.6504	False
1	5	-0.366	0.8256	-1.4598	0.7278	False
2	3	-0.0496	0.9992	-1.0215	0.9223	False
2	5	-0.1974	0.9732	-1.375	0.9802	False
3	5	-0.1478	0.9927	-1.5251	1.2295	False

- The F-statistic (0.587337) is relatively low, suggesting minimal variation in reimbursement ratios explained by racial differences.
- The p-value (0.623238) is greater than 0.05 (commonly used significance level). This indicates that we fail to reject the null hypothesis, which is that there's no statistically significant difference in reimbursement ratios based on race.
- There is no significant difference in the reimbursement ratios between different racial groups.

Visits and Claim Insights:

- Patients with chronic conditions are likely to have more frequent visits and higher total claim amounts due to ongoing medical needs.
- High variability in claim amounts and visits can indicate differences in healthcare needs and utilization patterns among patients.



Actionable Insights:

- Understanding the distribution of claim amounts can help in budgeting and forecasting healthcare costs.
- Identifying patients with exceptionally high costs can inform targeted interventions to manage their healthcare needs more efficiently.
- Understanding the racial distribution of Medicare beneficiaries in the context of overall U.S. demographics provides valuable insights into healthcare access, utilization, and disparities.