

Goal:
Find relationship between outpatient medicare reimbursement and the beneficiaries' chronic condition

Cleaned Data Summary Statistics

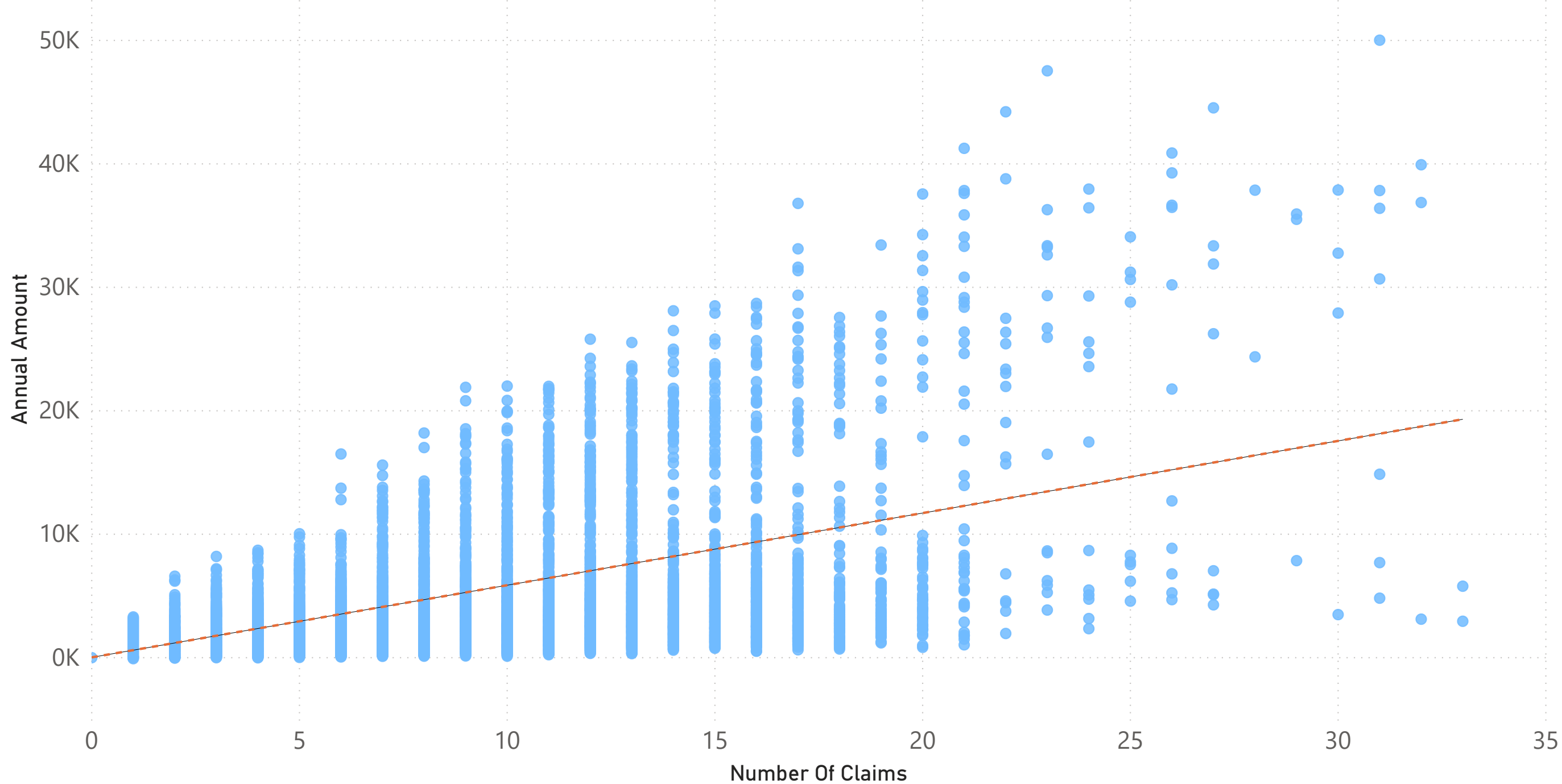
NumberOfClaims	Count	Max	Average	Median
Total	116352	33	2.42	1

Gender(1:M / 2:F)	Average Number Of Claims
2	2.54
1	2.27
Total	2.42

On average, Female has slightly higher claims than male.

Number Of Claims & Outpatient Medicare Reimbursement Annual Amount

They share positive relationship with 0.39 R Square



How To Define High Utilizer

▲ Num Of Beneficiaries With Total Claim <= 7

105347

▲ Num Of Beneficiaries With Total Claim > 14

1134

90th Percentile of *NumberOfClaims* is 7

99th Percentile of *NumberOfClaims* is 14

Thus, High Utilizers will be defined by the beneficiaries with more than 14 total claims

Diabetes(1:Y / 2:N)	Count of Beneficiaries	Sum of Super Utilizer ▲	Percentage of SuperUtilizer
2	72292	126	11.11%
1	44060	1008	88.89%
Total	116352	1134	100.00%

There is a high population(89%) in High Utilizers have diabetes.

The hospital can provide additional information about appropriateness of medical services and how to reach out to the primary healthcare providers. However, whether the High Utilizer use the outpatient service appropriately in our case needs further analysis.

Steps for this mini project

1. import the *Beneficiary* data and change the data type
2. import the *Outpatient* data and checked the data type
3. create new column *Year* based on the *CLM_THRU_DT* variable in *Outpatient* data
4. only *Year* of 2008 in *Outpatient* data will be used in this analysis
5. merge the *Beneficiary* and *Outpatient* data by counting the total number of outpatient claim for each beneficiary
6. use pivot table to find the summary statistics for the cleaned data
7. plot scatter plot to show the relationship between number of claims and annual medicare reimbursement amount
based on the scatter plot, the more claims tend to have higher reimbursement amount
8. use pivot table to find how to define the high utilizers
9. the max number of claims is 33 while the 99% of beneficiaries claims less than 14
10. find the common chronic condition among the high utilizer
11. about 89% of the high utilizers have diabetes

Reference

Using Claims Data. National Rural Health Resource Center. <https://www.ruralcenter.org/population-health-toolkit/data/using-claims-data>