1. Each row in the data represents a claim that was filed by a specific policy holder and contains information regarding the claim itself and the patient. Each patient (hashed\_ph\_id) can have several claims (hashed\_claim).

information about the claim- when was the claim filed, end date of the final decision, whether it was approved or denied

information about the patient- birthdate, age, gender, state where it was filed in the US, primary diagnosis of the patient, and whether the patient lives at his home or in any other settings.

1. missing values for claim\_end\_date are those where the claim\_status was denied, new or withdrawn. that means that claims that where denied don't have claim\_end\_date. new claims won't have end date as they are still looked at. and withdrawn cases won't have claim end date. The missing values are of type Missing not at random (MNAR)- (i.e., the value of the variable that's missing is related to the reason it's missing).

Non-missing values for eb\_reassessment\_date- those with status Approved-Reassessment and 1 withdrawn.

On the other hand, missing values occur in all the claim status categories. I tried to look for a trend whether the missing data occurs for specific claim status, gender filling state, last care setting, primary diagnosis, eb status but couldn't find a trend. that means that the missing data is of type missing completely at random (MCAR).

There is a peak of approved claims at around age 85. Claims before age 75 are rarely approved. Claims after age 93 are also less likely to be approved.

Chart, histogram

Description automatically generated

Age median of first time approved claimants=86.06849315068493

Age IQR of first time approved claimants=8.220547945205482

Males and females distribution of approved claims by age seems fairly similar. There are less cases of approved claims for men over the age of 90 but this might be due less claims filed.

Chart, histogram

Description automatically generated

1. There are less approved cases for males versus females upon first approved claim (not necessarily significant)

Chart, bar chart

Description automatically generated

Males filed less claims than females for the first claim (not necessarily significant)

Chart, bar chart

Description automatically generated

1. There are less denied claims for males vs females, but this graph is misleading as we need to look at the proportion of denied/approved by gender.

Chart, bar chart

Description automatically generated

When looking at the proportion of denied/approved in males and females we can see that the difference is not significant.

Chart, bar chart

Description automatically generated

z\_stat: -1.012, p\_value: 0.312

Fail to reject the null hypothesis - There is no significant difference in denying claims between genders