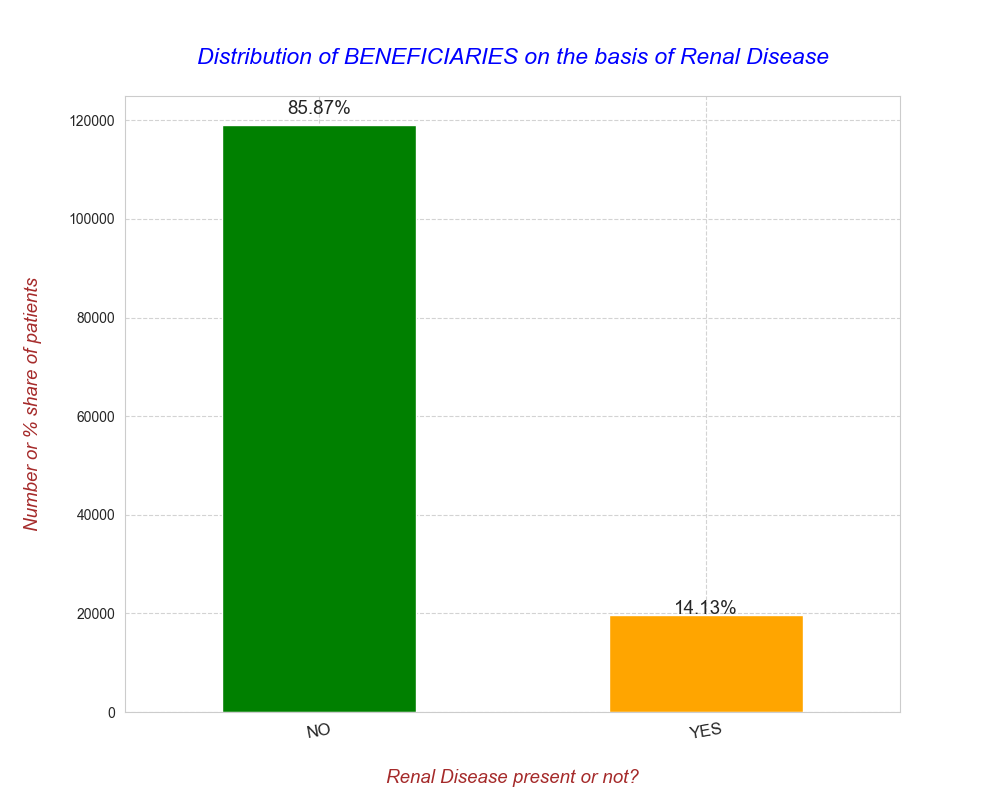
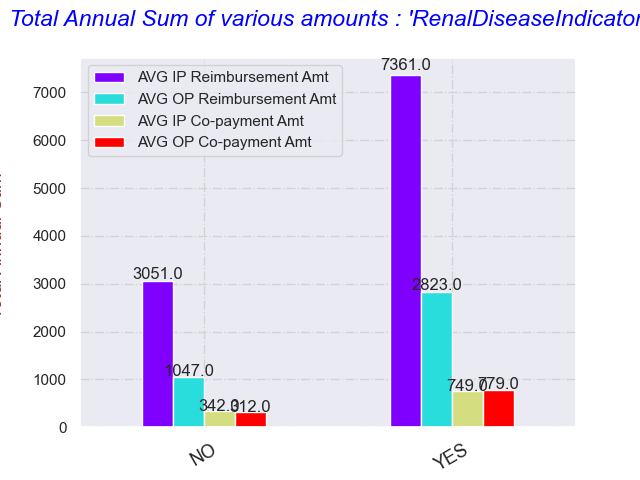
**Q1.How much is the number of beneficiaries with Chronic Renal Disease?**

* There are two main types of kidney disease - short-term (acute kidney injury) and lifelong (chronic).
* The two main types of kidney disease are short-term (acute kidney injury) and lifelong (chronic kidney disease).
  + Chronic kidney disease, also known as chronic renal disease or CKD, is a condition characterized by a gradual loss of kidney function over time.
* **What are the main causes of chronic kidney disease?**
  + Diabetes and high blood pressure, or hypertension, are responsible for two-thirds of chronic kidney disease cases.



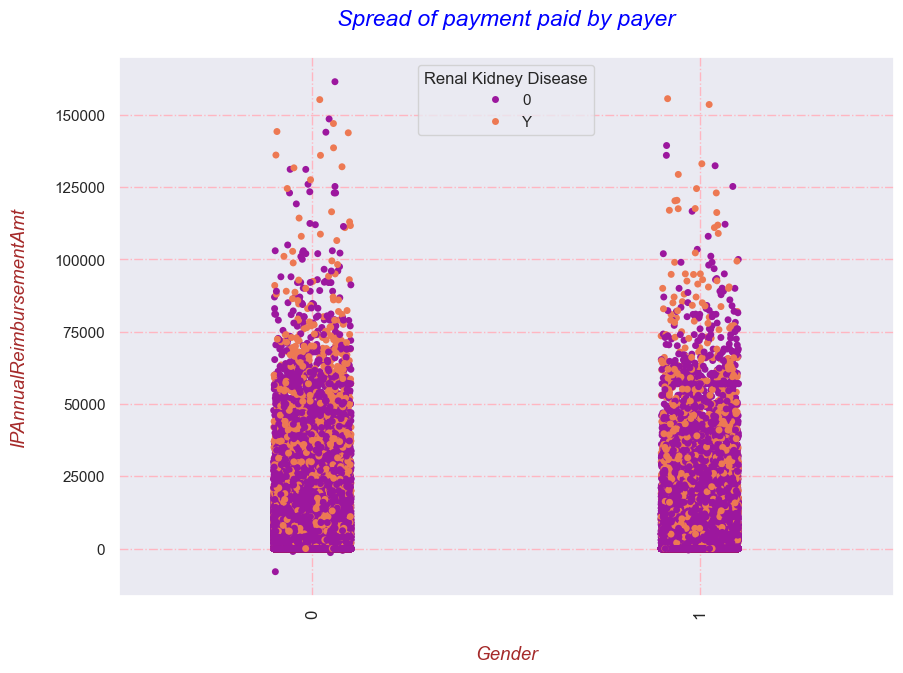
* **OBSERVATION**
  + Above graph tells us that around 14% of beneficiaries has or had Kidney Failure(Renal Disease).

### ****Q2. Let’s see the number of beneficiaries on the basis of 'RenalDiseaseIndicator'. And the Annual IP & OP expenditures for such patients.****



* **OBSERVATION**
  + The above graph is telling us below points:
    - Payer pays huge chunk of expenses specially when a beneficiary gets admitted with or without Renal Kidney Disease. To be more precise, in case of +ve the difference is more than 50%.
    - For other comparisons the difference is not very high.

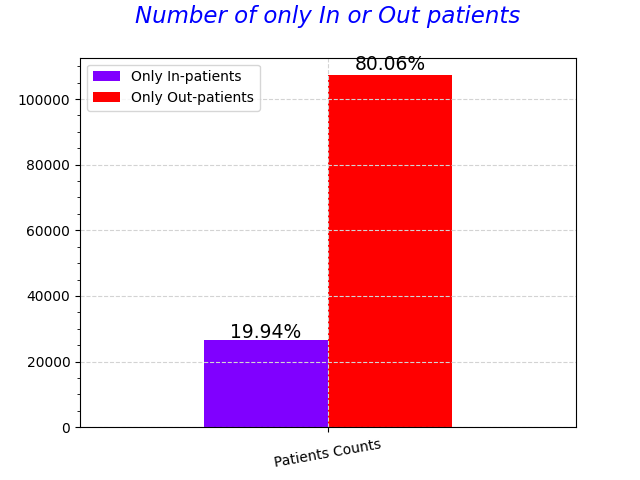
### ****Q2. Visualize the spread of pre-disease indicators for the Annual IP and OP expenditures across males and females.****



* **OBSERVATION**
  + The above graph shows us the complete overlapping of data points with some potential outliers(may be fraud).
    - Here, another thing that I found is that few of the points lying in the negative range(this is quite strange may be error).

# **IP & OP Data – EDA**

**Q1. Number of patients who either are in-patients or out-patients.**

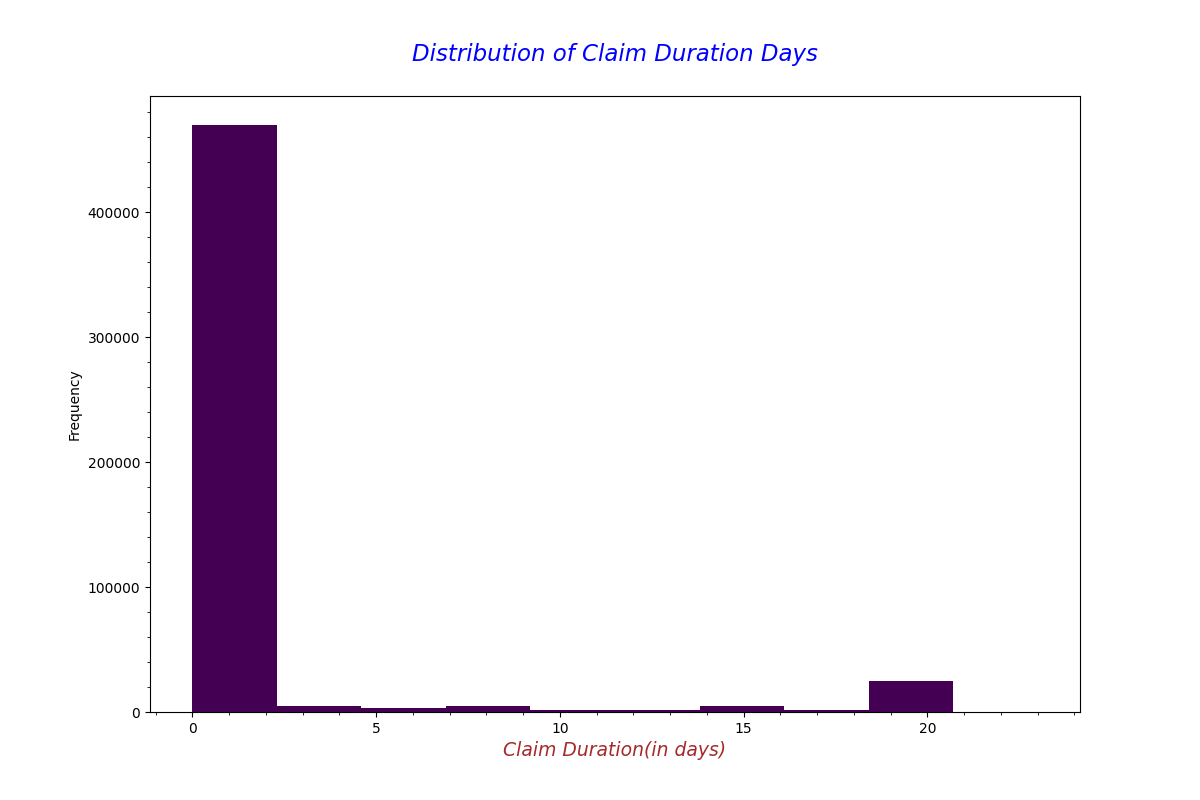


**OBSERVATION**

* From the above plot, we can decode that 80% of the patients gets medicated without even admission.

**Q2.How many days are the duration of Claim?**

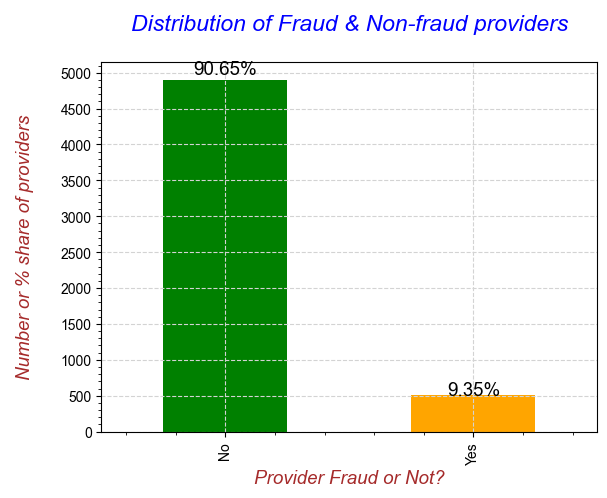
It is from 0 – 20 days.



**OBSERVATION**

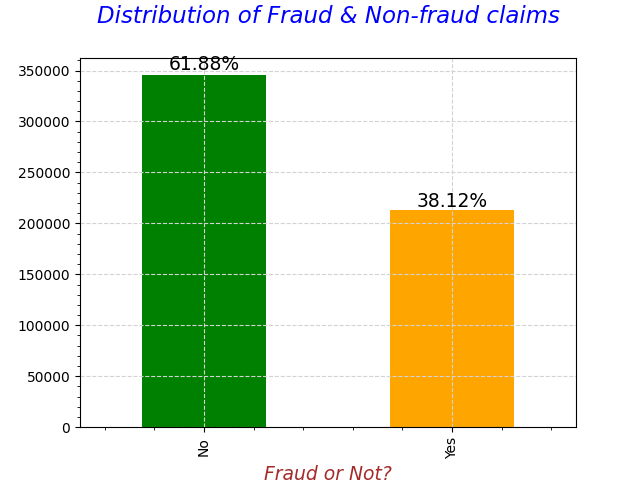
* From the above plot, we can decode that the majority of the claims filed for less than or equals to 2 days.

**Q3.Percentage distribution of Fraud and Non-Fraud** .



**OBSERVATION**

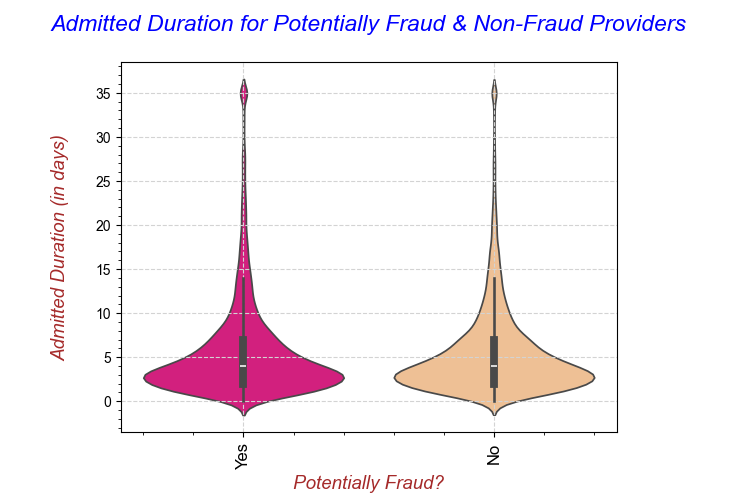
* From the above plot, we can say that 90% of the providers are not fraudsters and only 9% of them are involved in frauds.

**Q4. What’s the ratio of Fraud & Non-Fraud claims?**

**OBSERVATION**

* The above plot shows us that, 62% of claims are Non-Fraud and 32% of them are Fraudulent.

**Q5.Patient Admitted Duration for Potentially Fraud & Non-Fraud Providers.**

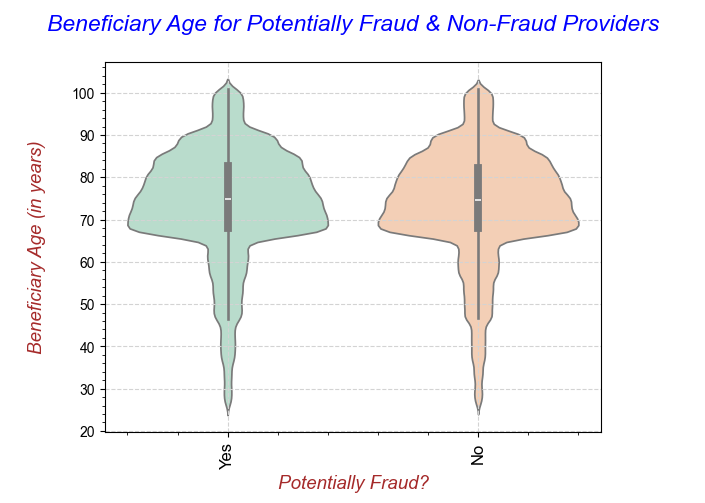
****

Maximum patients’ duration is around 0-5 days.

**OBSERVATION**

* The above plot clearly shows us that there is no difference in the distribution of Admit Duration for Potentially Fraud and Non-Fraud Providers.

**Q6.Finding out Beneficiary Age for Potentially Fraud & Non-Fraud Providers.**

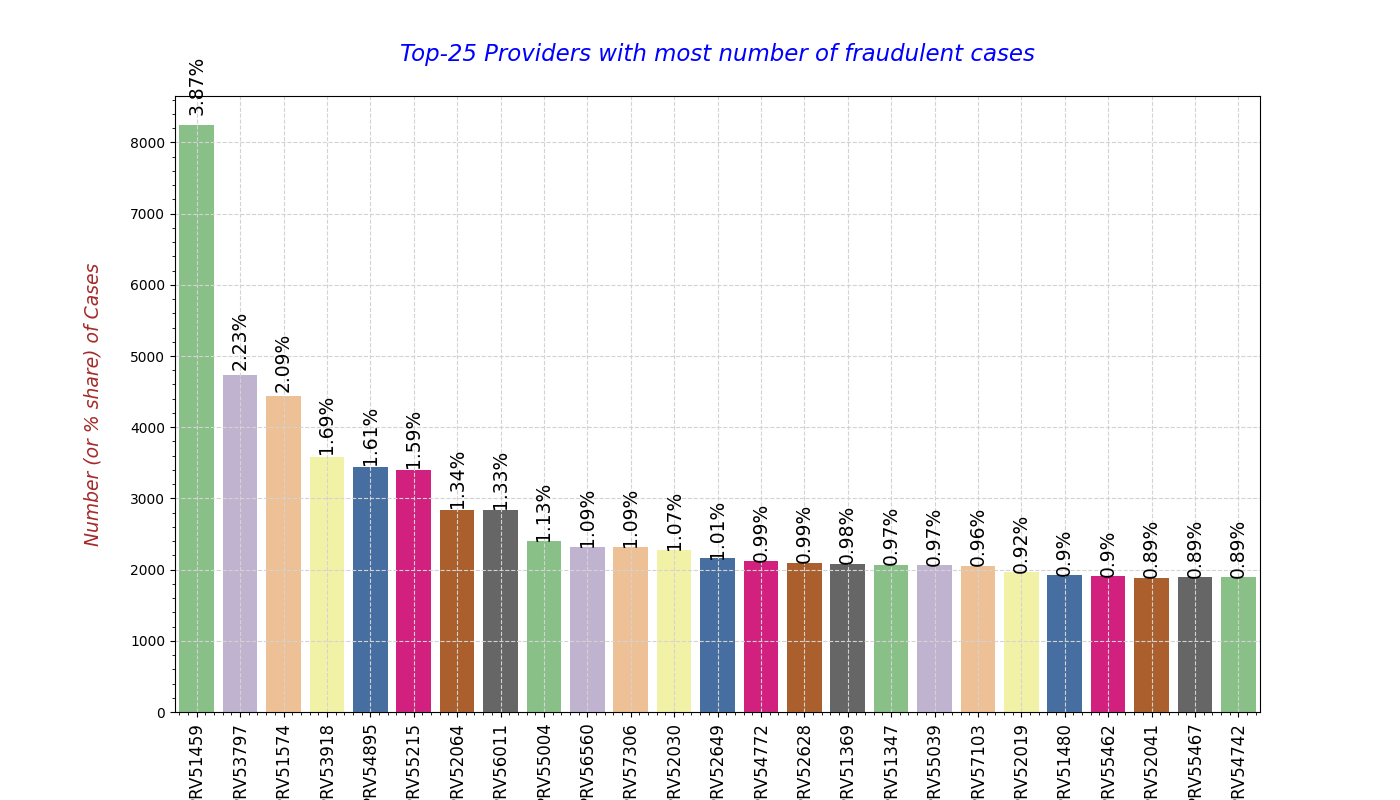
****

**From the above plot, the maximum beneficiaries can be noticed around the age 68 – 74 old years patients.**

**OBSERVATION**

* The above plot clearly shows us that there is no difference in the distribution of Beneficiary Age for Potentially Fraud and Non-Fraud Providers.

**Q.7 Which are the Top-25 Providers with maximum number of fraudulent cases?**

Q.7 Which are the Top-25 Providers with maximum number of fraudulent cases?

OBSERVATION

The above plot shows us the Top-25 Providers with most percentage of Fraudulent Case Submissions.

Here, PRV1459 has the highest percentage share of fraudulent cases. The difference between

Other providers are not that high.