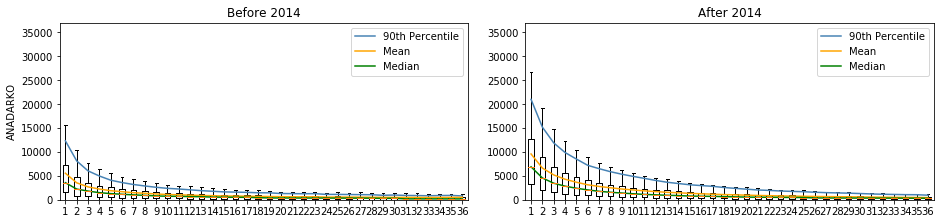
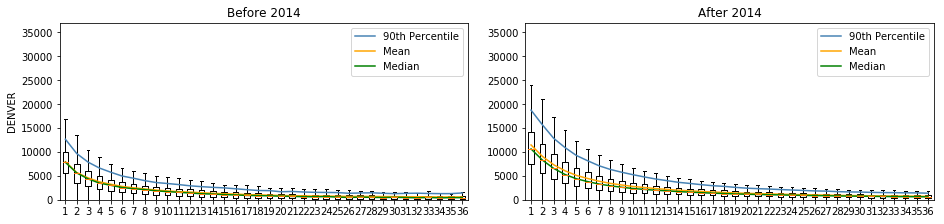
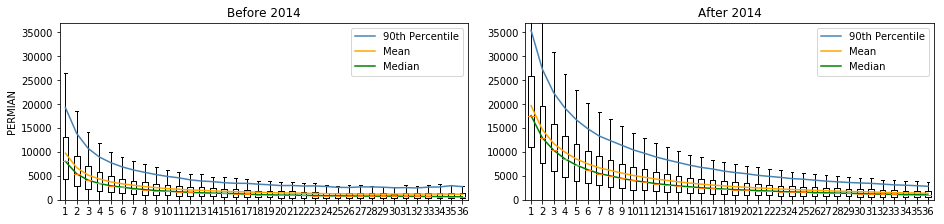
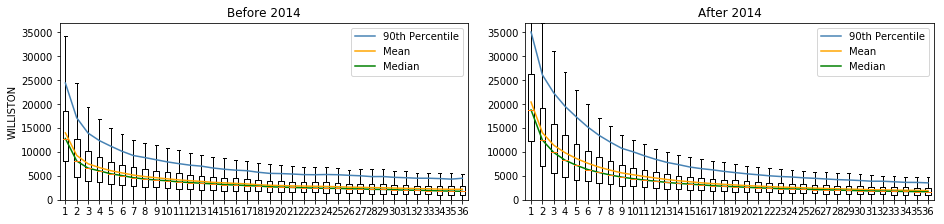
Insights:

**Dataset:**

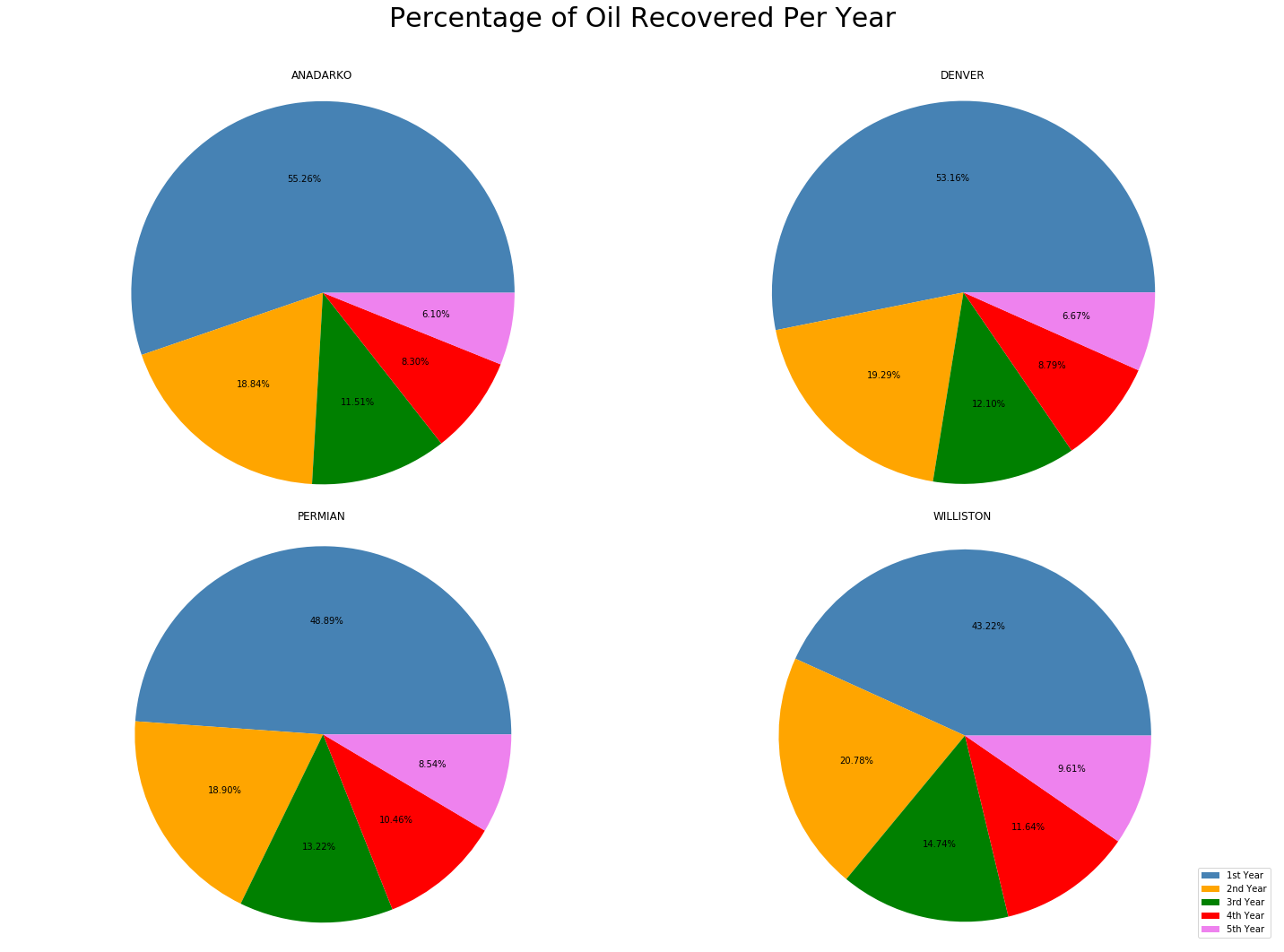
* Two of the APIs in the production\_data file are from the same basin and have the same operator. In addition to that they produced the exact amount of gas, oil, water at the same dates for all their life time. That’s why we believe that the 2 APIs refer to the same well
  + - 35073256530000
    - 35073256910000
* Some records were duplicated in the dataset.

**Plot 1:**

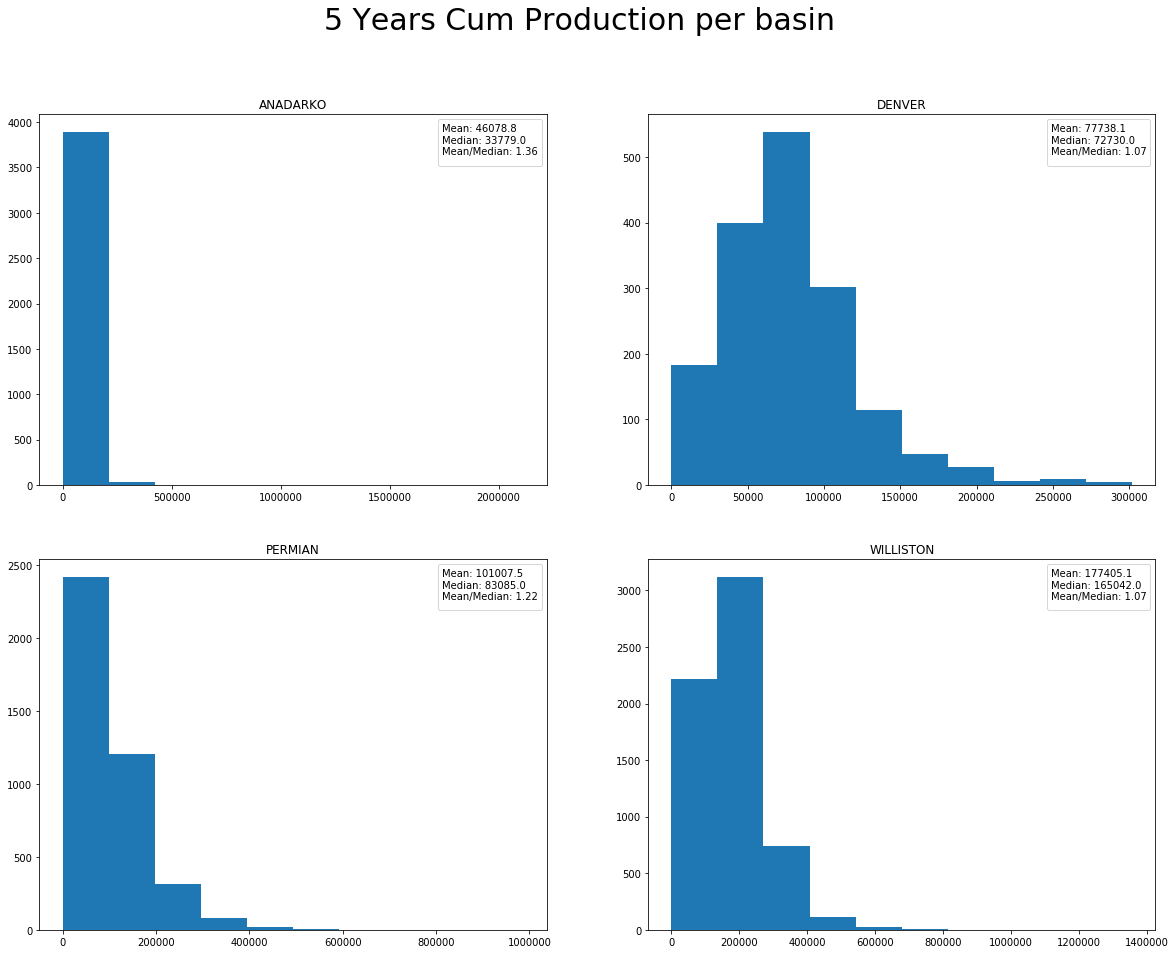
* The months at which the basins production attain a form of low production monotonic horizontal lines increased by a rough average of 11 months after applying the new drilling techniques in 2014.  
  **i.e.** around one year of higher levels of production is added to each basin with the new applied techniques.
* The production amount of wells in each basin acquires similar behavior resulting in close mean and median values. However, the difference between the mean and the median in “Anadarku” indicates the presence of outliers.

**Plot 2:**



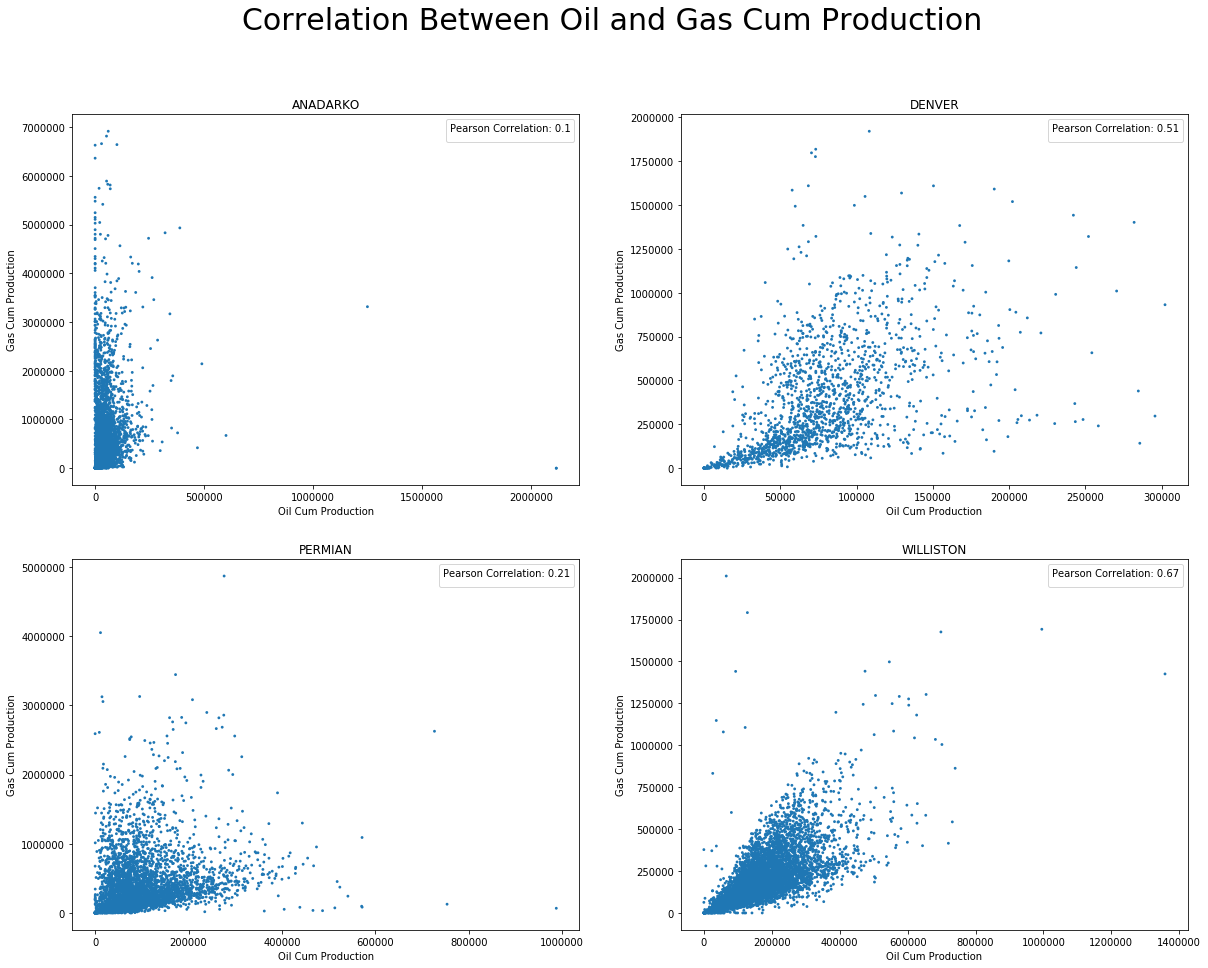
* With plot one taken into consideration, we can infer that basins starting with higher production volume in their peaks, decline with higher percentage rates than those starting with lower levels. **i.e.** peak production value is directly related to the production-declination percentage.

**Plot 3:**



* As the weight of the data gets centralized, the difference between the mean and the median values thus, it is obvious that “Anadarku Basin” has the highest number of outliers followed by “Permian Basin”.

**Plot 4:**



* Oil and gas production is always positively correlated. Moreover, the basins with outliers have smaller correlation coefficients. However, correlation doesn’t imply causation.

**Plot 5:**