Mortgage Pooling Challenge

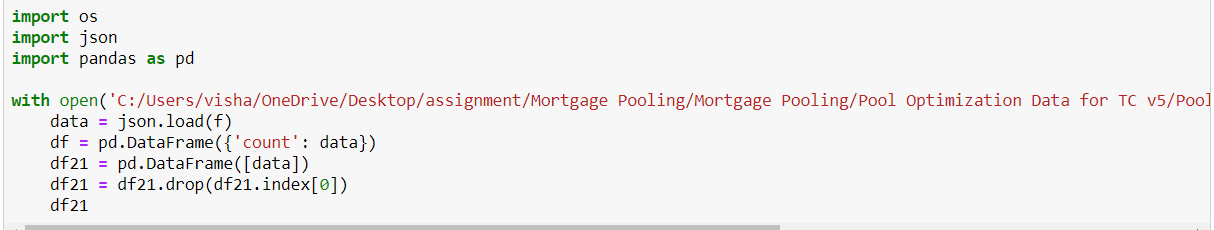
**Mission**

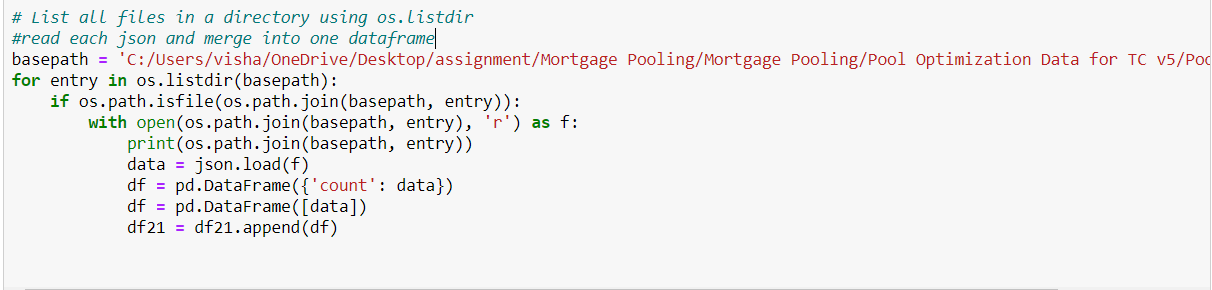
Given the background and the json dataset, what useful information would you predict? How

would you analyze the data? We want to see at least 5 questions being asked and answered

through your visualization tool.

* Data was consolidated using python script, all the json files were parsed to merge into a single data file to be used for analysis by BI tools.
* Please find below python code snippets for data consolidation







Visualization

* Data was studied to understand variable importance and definition
* Data was visualized using Tableau to analyze Mortgage pooling problem
* Questions used for analysis
  + What is the pool pricing associated with Loan Amount?
  + What is the pool pricing for Agency like Fannie Mae & Freddie Mac?
  + What is the average amount-based Fico values for High & Low Flag loans?
  + What is the breakdown of amount set for each Agency across different servicers?
  + How property type is being priced by each servicer?
  + State-wise spending on loan amount for loan purposes?
  + What is the average Debt-to-income ratio for each pool group?
* Tableau Visualizations have been published on tableau public profile as well as .twbx file has been uploaded on Github
* Useful Links:
  + Tableau profile: <https://public.tableau.com/profile/vishakh.s.narkar#!/>
  + GitHub profile: <https://github.com/vishakhnarkar/Mortgage-Pooling>