• AGENCY\_ID

• PRIMARY\_AGENCY\_ID – (contains missing values) master agency if part of group

• PROD\_ABBR – 33 products of which 14 are CL and 19 are PL

• PROD\_LINE – commercial lines (CL) or personal lines (PL)

• STATE\_ABBR

• STAT\_PROFILE\_DATE\_YEAR – data starts in mid-2005 and continues into 2015

• RETENTION\_POLY\_QTY – current number of policies that are still active from previous year

• POLY\_INFORCE\_QTY – number of policies active for that year

• PREV\_POLY\_INFORCE\_QTY – (contains missing values) number of policies active in the previous year

• NB\_WRTN\_PREM\_AMT – new business in written premium

• WRTN\_PREM\_AMT – total written premium

• PREV\_WRTN\_PREM\_AMT – (contains missing values) written premium during the same period in the previous year

• PRD\_ERND\_PREM\_AMT – amount of premium taken in • PRD\_INCRD\_LOSSES\_AMT – losses • MONTHS – number of months included in the data for that year; the original data was monthly and some months were missing so the aggregate doesn’t make up the entire year if the months value is less than 12

• RETENTION\_RATIO – (computed & contains missing values) computed for each row in the data as RETENTION\_POLY\_QTY / PREV\_POLY\_INFORCE\_QTY; therefore it’s a granular measure of the retention for that agency writing that particular product in that particular state from the previous year

• LOSS\_RATIO – (computed & contains missing values) computed for each row in the data as PRD\_INCRD\_LOSSES\_AMT / WRTN\_PREM\_AMT; currently I’m only computing results where the WRTN\_PREM\_AMT is greater than 0; however there are many cases where there are losses but no premium maybe from a previous claim that’s still being paid, so I included codes to indicate whether there are positive or negative losses on zero premiums; they are located at the end of this document

• LOSS\_RATIO\_3YR – (computed & contains missing values) computed by agency by line of business by year for the three year period ending in that year, if there is data for three years, otherwise the two years or one year of data available; to make this more tangible, the first complete year of data we have for an agency will have the loss ratio for that year; the second complete year of data will have the mean loss ratio for that year and the previous year; the third complete year of data will have the mean loss ratio for those three years; then the fourth and greater will have the mean of the three year period ending in that year; note that the mean loss ratios are computed independently for PL and CL

• GROWTH\_RATE\_3YR – (computed & contains missing values) computed by agency by line of business by year for the three year period ending in that year; measures the average growth in written premium for that agency in that line of business; only computes results for agencies that have data for the entire range of years; since the measure is over three years of growth, there needs to be a base year to be used as a standard so four years of data are needed; in order to include as many results as possible, the PREV\_WRTN\_PREM\_AMT column is used if it exists in the first year of data available, so that it can be used as a base year, otherwise the WRTN\_PREM\_AMT is the only column used

• AGENCY\_APPOINTMENT\_YEAR – (contains missing values) year the agency started doing business with Azure

• ACTIVE\_PRODUCERS – (contains missing values) number of active producers in the agency • MAX\_AGE – (contains missing values & results may not be accurate) maximum age producer at that agency

• MIN\_AGE – (contains missing values) minimum age producer at that agency

• VENDOR\_IND – indicator column to specify whether the agency subscribes to a vendor

• VENDOR – (contains missing values) the vendor that the agency subscribes to • PL\_START\_YEAR – (contains missing values) year the agency started using the PL vendor

• PL\_END\_YEAR – (contains missing values) year the agency stopped using the PL vendor

• COMMISIONS\_START\_YEAR – (contains missing values) year the agency started using the COMMISIONS vendor

• COMMISIONS\_END\_YEAR – (contains missing values) year the agency stopped using the COMMISIONS vendor

• CL\_START\_YEAR – (contains missing values) year the agency started using the CL vendor

• CL\_END\_YEAR – (contains missing values) year the agency stopped using the CL vendor

• ACTIVITY\_NOTES\_START\_YEAR – (contains missing values) year the agency started using the ACTIVITY NOTES vendor

• ACTIVITY\_NOTES\_END\_YEAR – (contains missing values) year the agency stopped using the ACTIVITY NOTES vendor

• CL\_BOUND\_CT\_MDS – (contains missing values) number of bound policies quoted through a MDS (probably a data recording error, should be DSM) in the current year to date, that is the first six months of 2015, in commercial lines

• CL\_QUO\_CT\_MDS – (contains missing values) number of quoted policies through a MDS (probably a data recording error, should be DSM) in the current year to date, that is the first six months of 2015, in commercial lines

• CL\_BOUND\_CT\_SBZ – (contains missing values) number of bound policies quoted through a SBZ in the current year to date, that is the first six months of 2015, in commercial lines

• CL\_QUO\_CT\_SBZ – (contains missing values) number of quoted policies through a SBZ in the current year to date, that is the first six months of 2015, in commercial lines

• CL\_BOUND\_CT\_eQT – (contains missing values) number of bound policies quoted through an eQT in the current year to date, that is the first six months of 2015, in commercial lines

• CL\_QUO\_CT\_eQT – (contains missing values) number of quoted policies though an eQT in the current year to date, that is the first six months of 2015, in commercial lines

• PL\_BOUND\_CT\_ELINKS – (contains missing values) number of bound policies quoted through ELINKS since September 2013 in personal lines

• PL\_QUO\_CT\_ELINKS – (contains missing values) number of quoted policies though ELINKS since September 2013 in personal lines

• PL\_BOUND\_CT\_PLRANK – (contains missing values) number of bound policies quoted through PLRANK since September 2013 in personal lines

• PL\_QUO\_CT\_PLRANK – (contains missing values) number of quoted policies though PLRANK since September 2013 in personal lines

• PL\_BOUND\_CT\_eQTte – (contains missing values) number of bound policies quoted through eQTte since September 2013 in personal lines

• PL\_QUO\_CT\_eQTte – (contains missing values) number of quoted policies though eQTte since September 2013 in personal lines

• PL\_BOUND\_CT\_APPLIED – (contains missing values) number of bound policies quoted through APPLIED since September 2013 in personal lines • PL\_QUO\_CT\_APPLIED – (contains missing values) number of quoted policies though APPLIED since September 2013 in personal lines

• PL\_BOUND\_CT\_TRANSACTNOW – (contains missing values) number of bound policies quoted through TRANSACTNOW since September 2013 in personal lines

• PL\_QUO\_CT\_TRANSACTNOW – (contains missing values) number of quoted policies though TRANSACTNOW since September 2013 in personal lines