**Metadata for calculated metrics and dataframes**

**Pandas Dataframes:**

**zllw\_df**: Dataframe containing entire dataset from Zillow CSV file

**zllw\_NY\_df**: Dataframe containing only NY state zipcode data from Zillow file

**zllw\_df\_cleansed**: Cleansed Zillow dataset with only NY state data and median property price for latest 60 months (5 years)

**abnb\_df**: Dataframe containing entire dataset from Airbnb CSV file

**abnb\_NY\_df:** Dataframe containing Airbnb data filtered on NY listings and only 2 bedroom properties with only columns that are required for analysis.

**abnb\_zllw\_df** : Dataframe created by left joining zllw\_df\_cleansed with abnb\_NY\_df

**abnb\_unmatched:**  Dataframe with airbnb data that did not find a matching zipcode in Zillow data

**city\_5year\_median\_property\_price:** a Rolled up dataframe at City level with a median property value at City level

**abnb\_match\_pass1 & abnb\_match\_pass2**: Final Merged dataframes.

**Calculated Metrics:**

Zllw\_df\_cleansed

|  |  |
| --- | --- |
| **Field** | **Description** |
| Current\_price | Most recent median price available in the dataset (as of 2017-06) |
| 1year\_Median\_Price | Calculated Median of most recent 12 months’ property rate (2016-07 to 2017-06) |
| 3year\_Median\_Price | Calculated Median of most recent 36 months’ property rate (2014-07 to 2017-06) |
| 5year\_Median\_Price | Calculated Median of most recent 60 months’ property rate (2012-07 to 2017-06) |

city\_5year\_median\_property\_price

|  |  |
| --- | --- |
| **Field** | **Description** |
| 5year\_Median\_Price | Calculated median property value at City level based on 5 year (60 months) median monthly property value |

abnb\_match\_pass1

|  |  |
| --- | --- |
| **Field** | **Description** |
| Rate\_of\_ROI\_5year | Calculated based on the formula  Listing price X 365 X 0.75/5year\_Median\_Price\*  *\*5year\_Median\_Price is from ‘zllw\_df\_cleansed’ dataframe* |

abnb\_match\_pass2

|  |  |
| --- | --- |
| **Field** | **Description** |
| 5year\_Median\_Price\_y | Calculated Median of matched zipcodes in the given city in |
| Rate\_of\_ROI\_5year | Calculated based on the formula  Listing price X 365 X 0.75/5year\_Median\_Price\_y\*\*  *\*\* 5year\_Median\_Price is from ‘city\_5year\_median\_property\_price’ dataframe* |