

[Weekly Report]

- Report No.1
- Date: Nov. 01, 2023~ Nov. 07, 2023
- Team: Yang Junseob (2019034639), Ryu Seung Gwon (2019087147)

		Highlights	Self Evaluation
Last Week	Summary	<ul style="list-style-type: none"> • Overview of Team Project 	Medium
Current Week	Baseline Goals (Given)	<ul style="list-style-type: none"> • Importing data in R • conducting Exploratory Data Analysis to examine key features of each file 	Medium
	Additional Goals (O.Y.O)	<ul style="list-style-type: none"> • Visualize data for understanding the statistical data distributions 	Medium
	Key Issues to Be Resolved	<ul style="list-style-type: none"> • Several duplicate data included in geo1 and geo3 datasets. • Some missing values included in LATITUDE_NBR, LONGITUDE_NBR, ASSESSOR_FINAL_ADDR_LN, ASSESSOR_CITY_NM columns. • Identify key characteristic of variables 	Medium
	Strategies	<ul style="list-style-type: none"> • Step 1: Conduct EDA for Analytics • Step 2: Perform descriptive statistics and data visualization. • Step 3: Identify any patterns, trends, duplicate, missing values and cleaning them 	Medium
	Results	<ul style="list-style-type: none"> • gain insights into our data and identify patterns, relationships. • Merge duplicate values and delete missing values (geo1 : 153021 obs -> 77101 obs, geo 3 : 1210596 obs -> 638270 obs) 	Medium
	Implications	<ul style="list-style-type: none"> • It helps us prepare the data for further analysis or modeling. 	Medium
Next Week	Things to Do	<ul style="list-style-type: none"> • Data Preparation • Exploring data from American Community Survey • Select Model for Analysis 	Medium
Remarks	Core Libraries & Packages	<ul style="list-style-type: none"> • Data Cleansing: dplyr • Visualization: ggplot2 	Medium
	Additional Remarks		

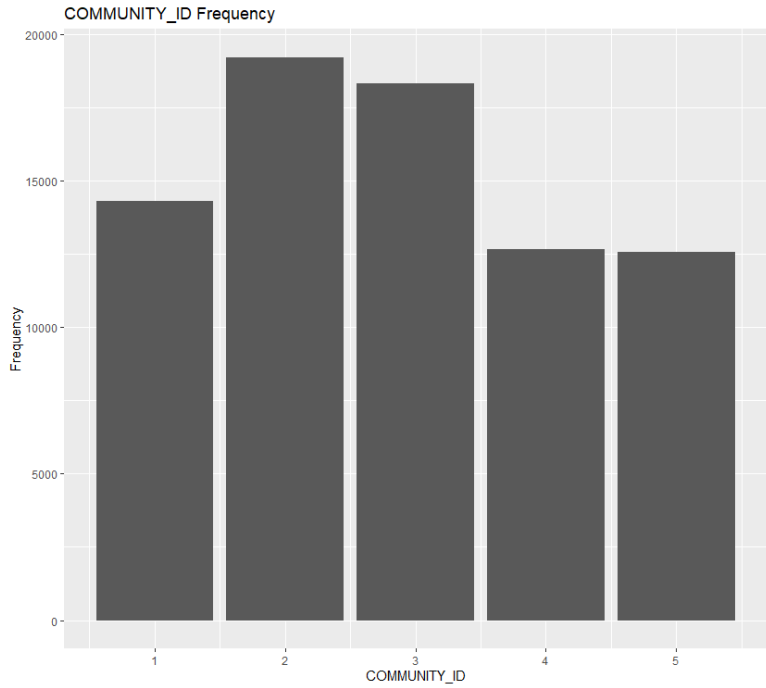
- Remarks
- Visualizations are given in Appendix.

- Work Progress

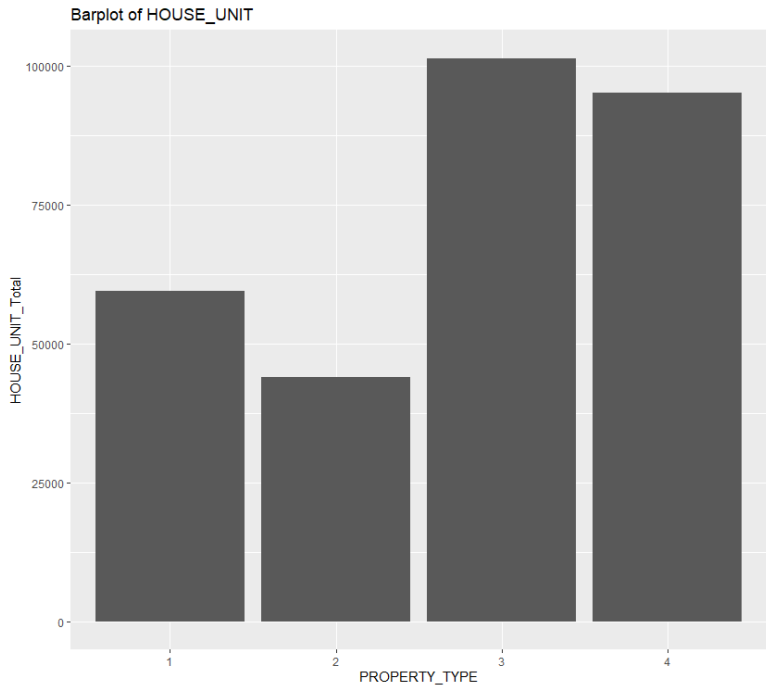
	Week 1	Week 2	Week 3	Week 4	Week 5
Yang Junseob	EDA	-	-	-	-
Ryu Seung Gwon	EDA	-	-	-	-

	Week 6	Week 7	Week 8	Week 9	Week 10
Yang Junseob	-	-	-	-	Final Report
Ryu Seung Gwon	-	-	-	-	Final Report

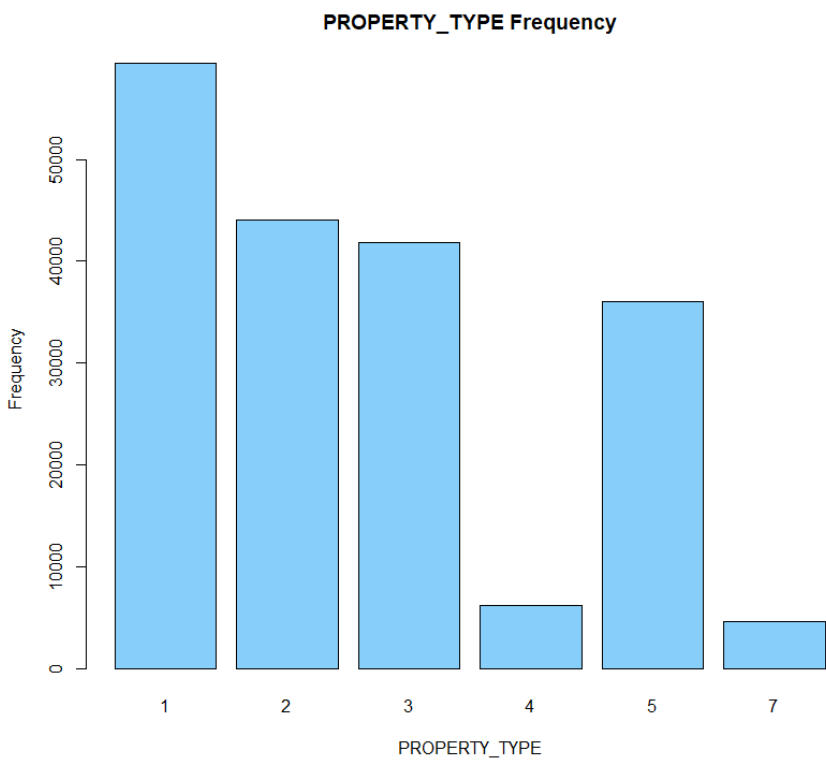
[Appendix1. Frequency of Community ID]



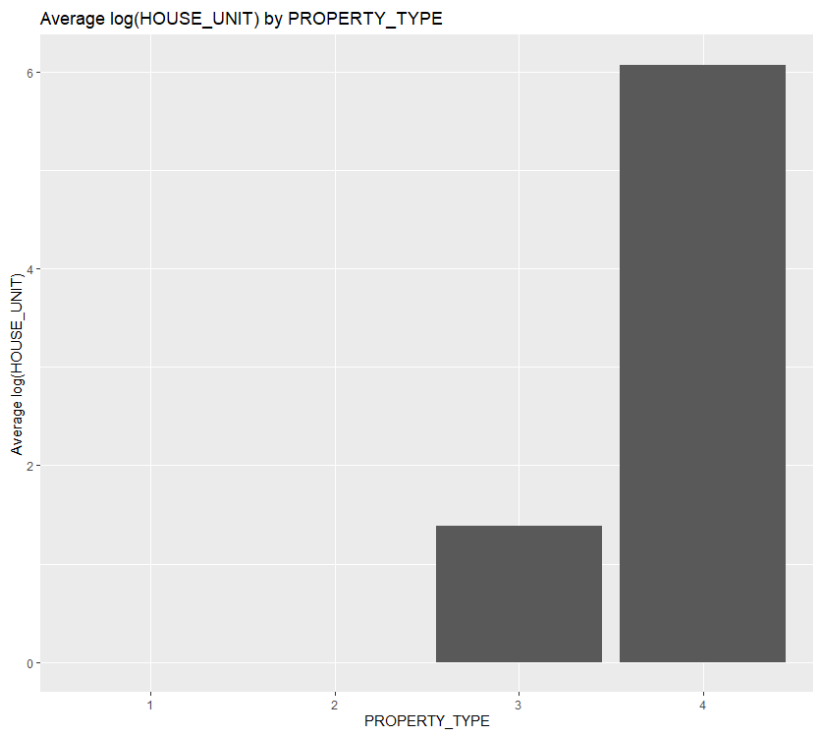
[Appendix2. Total House unit by Property type of house]



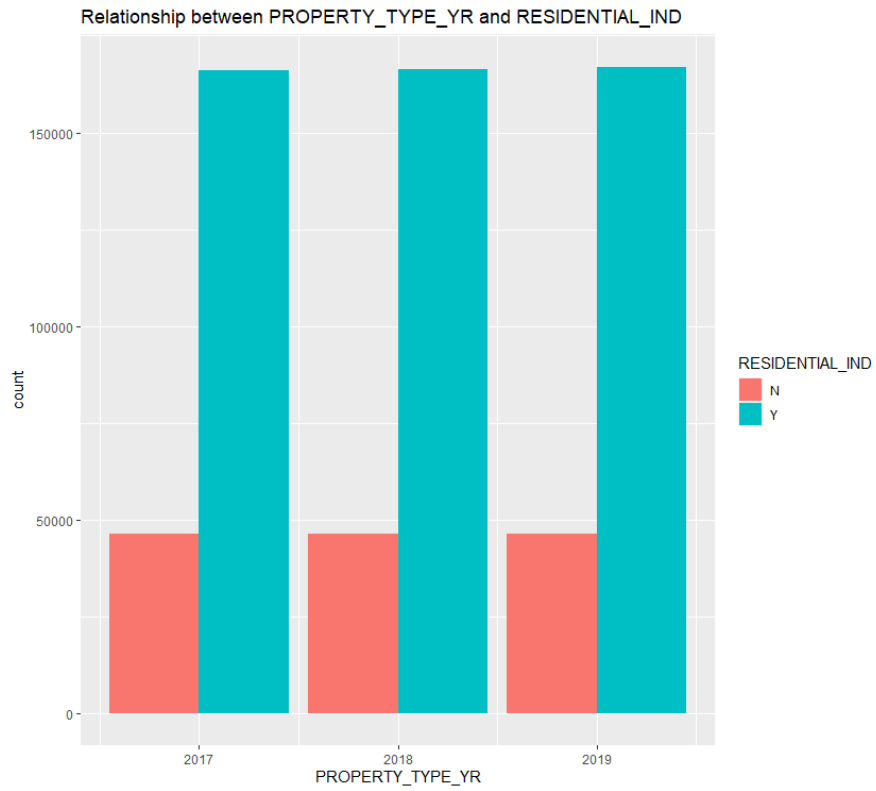
[Appendix3. Frequency of Property type]



[Appendix4. Average log units by property type of house]



[Appendix5. Relation between Property type and Residential Id by years]



[Appendix6. Distribution of assessor city in “geo3”]

