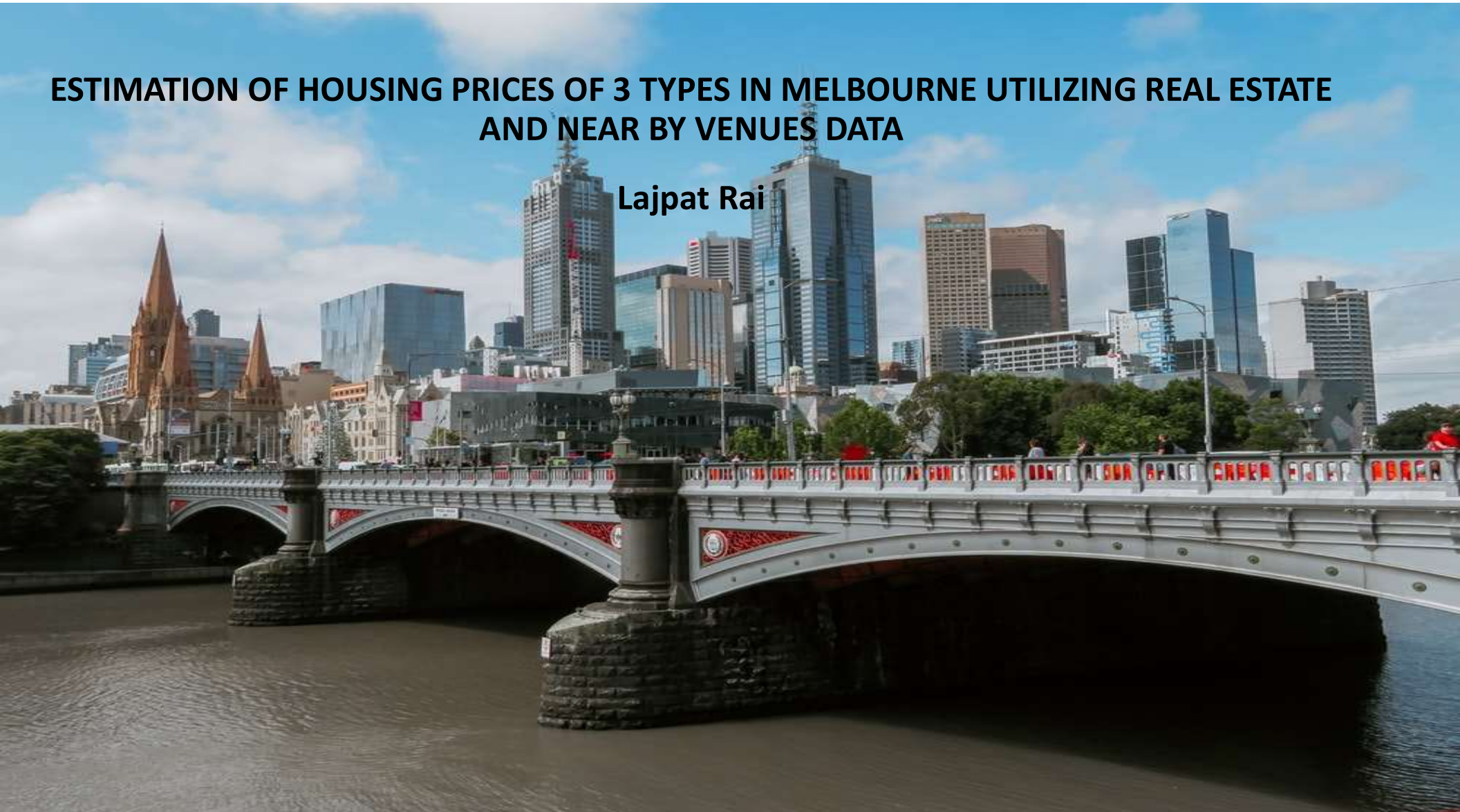


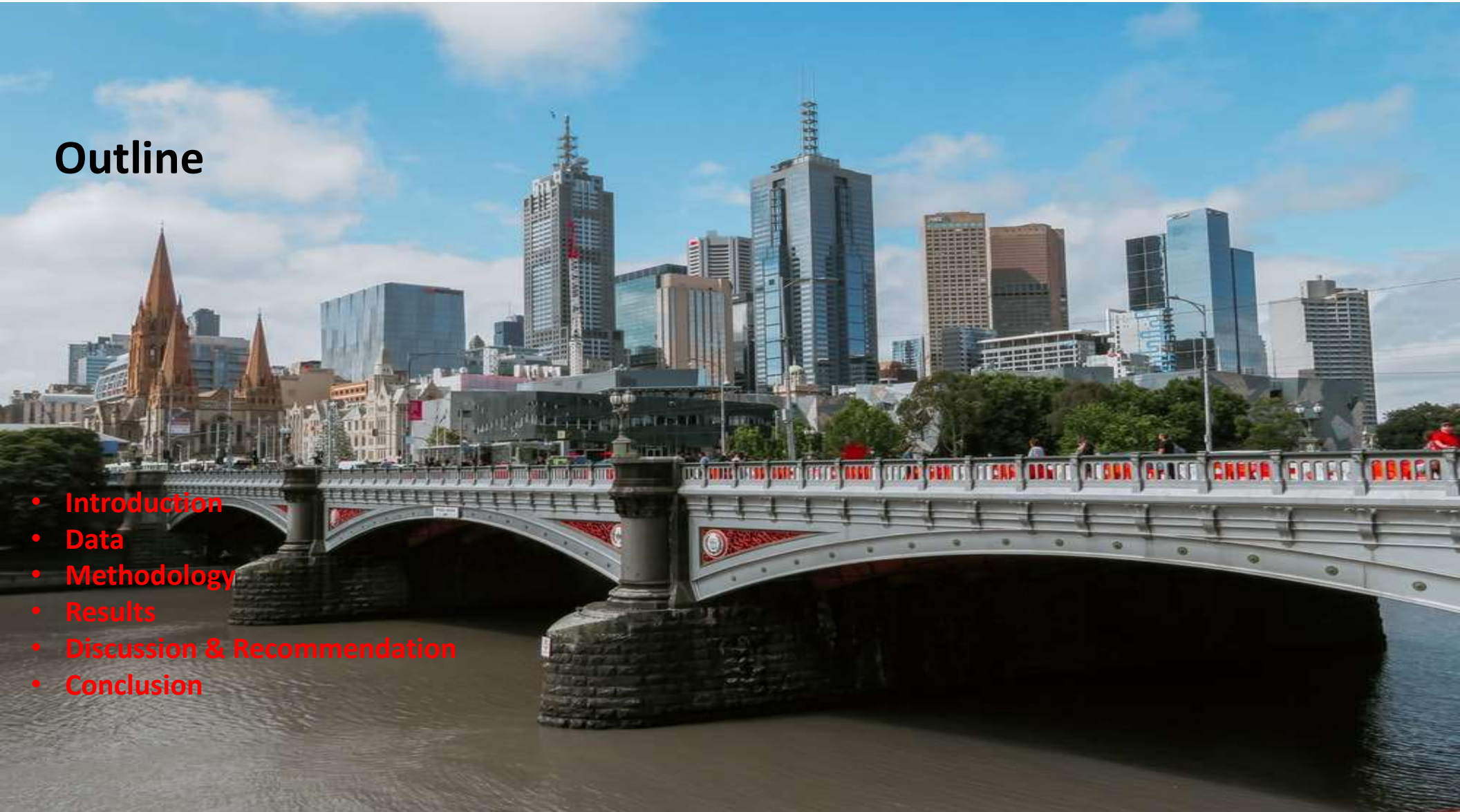
# ESTIMATION OF HOUSING PRICES OF 3 TYPES IN MELBOURNE UTILIZING REAL ESTATE AND NEAR BY VENUES DATA

Lajpat Rai



# Outline

- Introduction
- Data
- Methodology
- Results
- Discussion & Recommendation
- Conclusion





# Introduction

## Background

- Melbourne is the capital and most populous city of the Australian state of Victoria with enormous employment & business opportunities
- These recompenses have gained attention of several local Australian and International Immigrants.
- But with all these amenities, there lies Melbourne's real estate market hype



## Business Problem

- Challenge in selecting suitable property type for local & internationals to move in Melbourne due to its hot estate market.
- Affordability, accommodation features and nearby facilities always remain qualifying parameters for selection of an appropriate place to live for them
- The project seeks to explore real estate data to get an insight of property price variation in combination with its traits and near services

# Data

### Data Source # 1

Real estate data was taken from Kaggle.com

	Suburb	Address	Rooms	Type	Price	Method	SellerG	Date	Distance	Postcode	...	Bathroom	Car	Landsize	BuildingArea	YearBuilt
0	Abbotsford	68 Studley St	2	h	NaN	SS	Jellis	2016-03-09	2.5	3067.0	...	1.0	1.0	126.0	NaN	NaN
1	Abbotsford	85 Turner St	2	h	1480000.0	S	Biggin	2016-03-12	2.5	3067.0	...	1.0	1.0	202.0	NaN	NaN
2	Abbotsford	25 Bloomburg St	2	h	1035000.0	S	Biggin	2016-04-02	2.5	3067.0	...	1.0	0.0	156.0	79.0	1900.0
3	Abbotsford	18/659 Victoria St	3	u	NaN	VB	Rounds	2016-04-02	2.5	3067.0	...	2.0	1.0	0.0	NaN	NaN
4	Abbotsford	5 Charles St	3	h	1465000.0	SP	Biggin	2017-04-03	2.5	3067.0	...	2.0	0.0	134.0	150.0	1900.0

### Data Source # 2

Neighborhood venue category data was extracted using API Foursquare

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Abbotsford	-37.80385	144.996577	Three Bags Full	-37.807318	144.996603	Café
1	Abbotsford	-37.80385	144.996577	The Kitchen at Weylands	-37.805311	144.997345	Café
2	Abbotsford	-37.80385	144.996577	The Park Hotel	-37.802769	144.997029	Pub
3	Abbotsford	-37.80385	144.996577	Stomping Ground Brewery & Beer Hall	-37.804683	144.991171	Brewery
4	Abbotsford	-37.80385	144.996577	Laird Hotel	-37.805309	144.993124	Gay Bar



# Methodology

## Exploratory Data Analysis

- Histogram
- Pairplot
- Heat map
- Box plot
- Line plot



## Model Development & Evaluation: Machine Learning Approach

- Linear regression
- Multilinear regression
- Polynomial regression
- Ridge Regression
- Lasso Regression
- Gradient Booster regressor
- Random Forest Regression

# Exploratory Data Analysis

- **Histogram**

- Histogram for all property type was generated in order to provide an insight of the pricing frequency in the city of Melbourne in general which lies between \$ 800,000 - 120,000,0.

- **Pairplot**

- **Heat map**

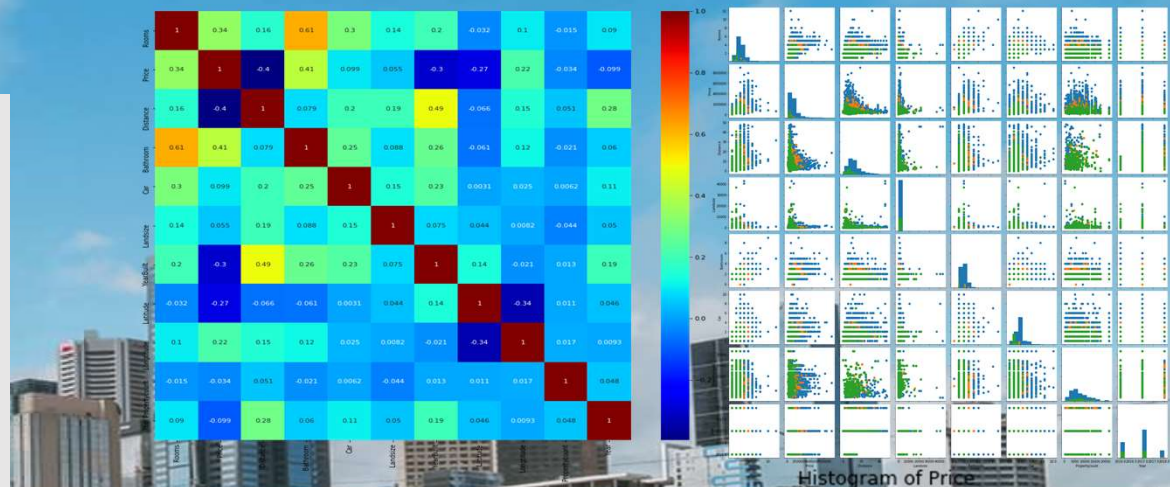
- Heat map & Pair plot suggest the strong to moderate relationship of price with number of rooms, distance from CBD, number of bathrooms & year built.

- **Box plot**

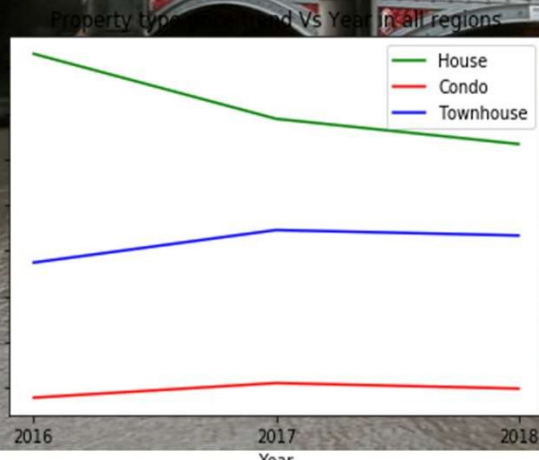
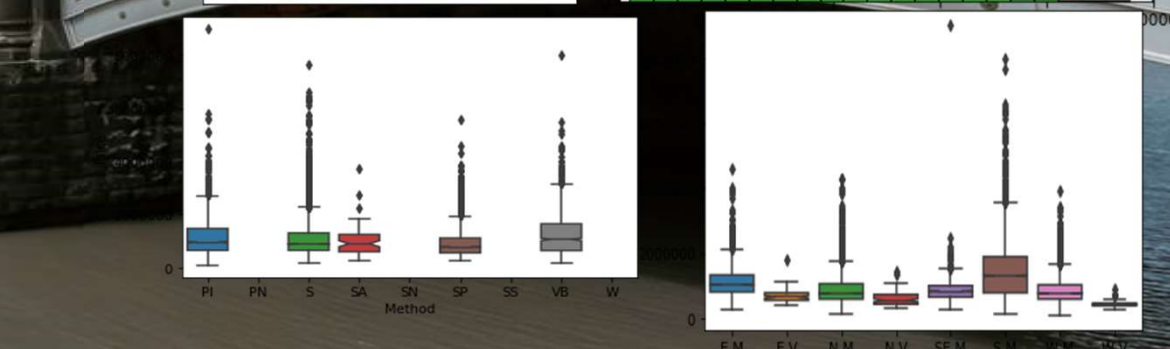
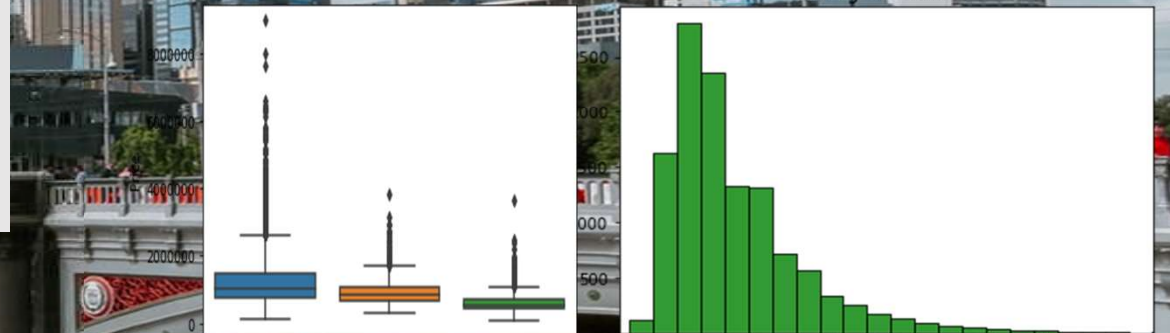
- Box plot was generated to identify the relationship between price & categorical variables.

- **Line plot**

- Line plot between mean price of all three types of properties over the period of 3 years in Melbourne city



Histogram of Price

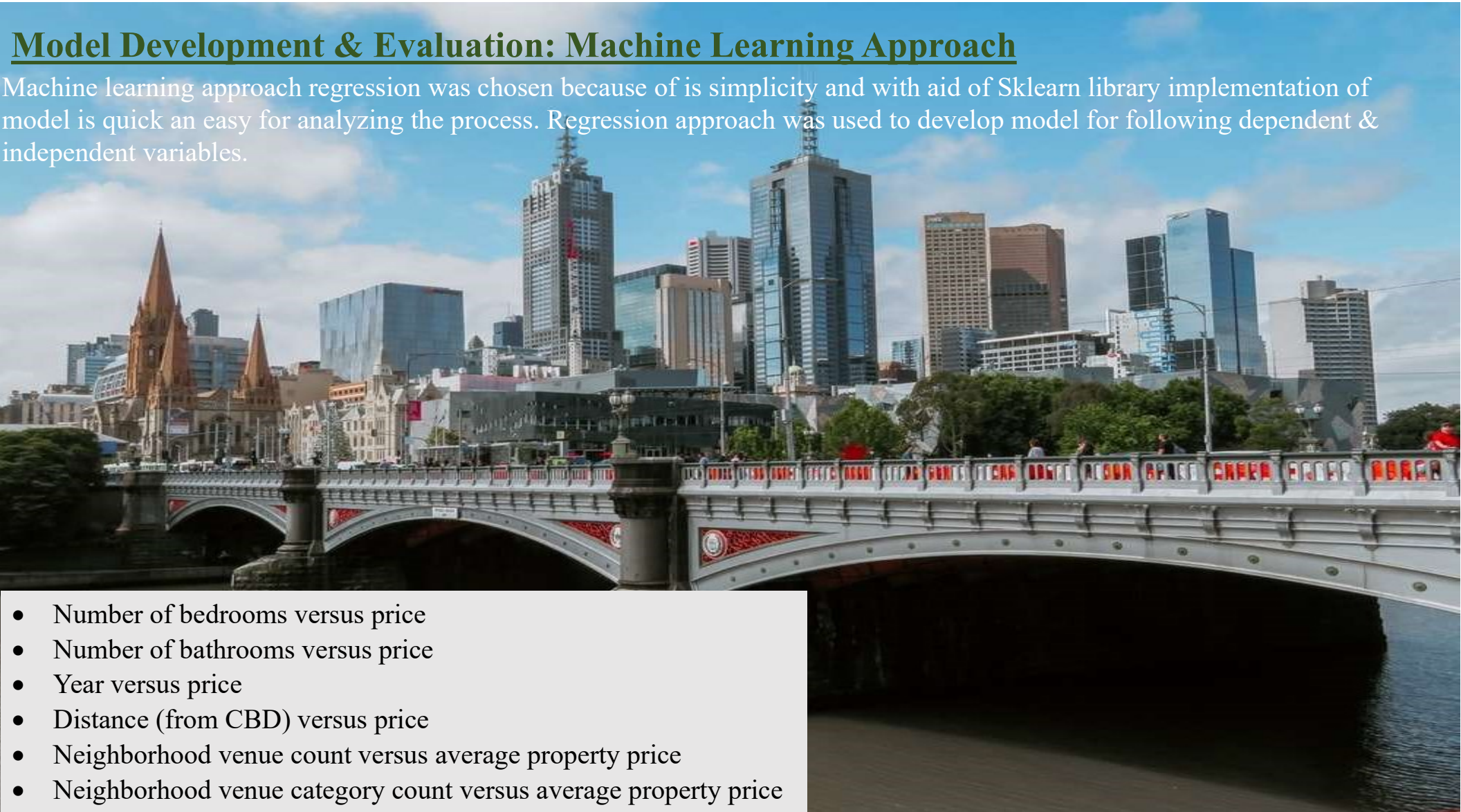




## **Model Development & Evaluation: Machine Learning Approach**

Machine learning approach regression was chosen because of its simplicity and with aid of Sklearn library implementation of model is quick and easy for analyzing the process. Regression approach was used to develop model for following dependent & independent variables.

- Number of bedrooms versus price
- Number of bathrooms versus price
- Year versus price
- Distance (from CBD) versus price
- Neighborhood venue count versus average property price
- Neighborhood venue category count versus average property price



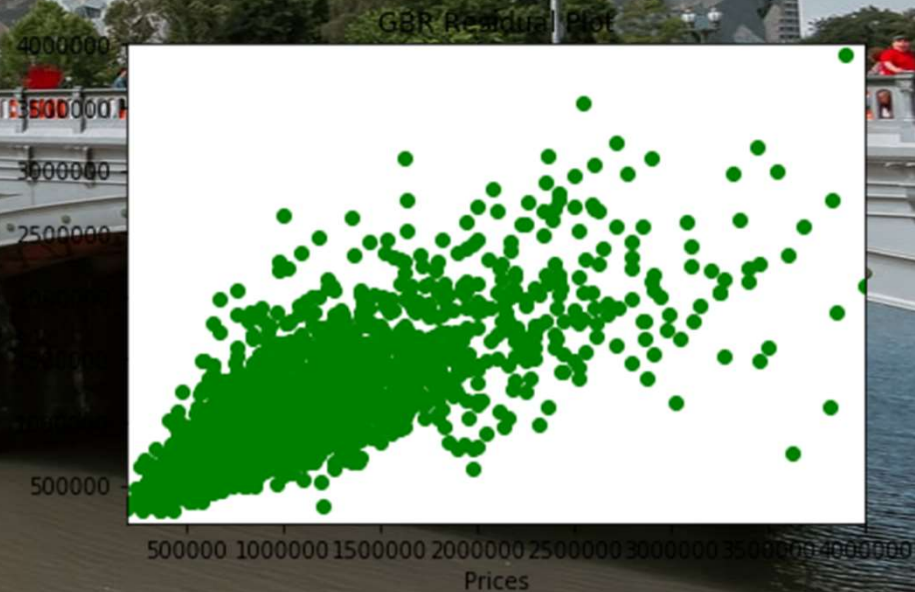
## Multilinear Regression

Graph below suggests that fitted values are reasonably close to the actual values, since the two distributions overlap a bit



## Gradient Booster Regression

GBR resulted in regression coefficient ( $R^2$ ) of 0.62 and mean squared error (MSE) of  $1.587E+11$  as linear relationship between predicted price versus actual price can be seen in graph below





## Comparison of all regression models for all type of properties in Southern Metropolitan region of Melbourne



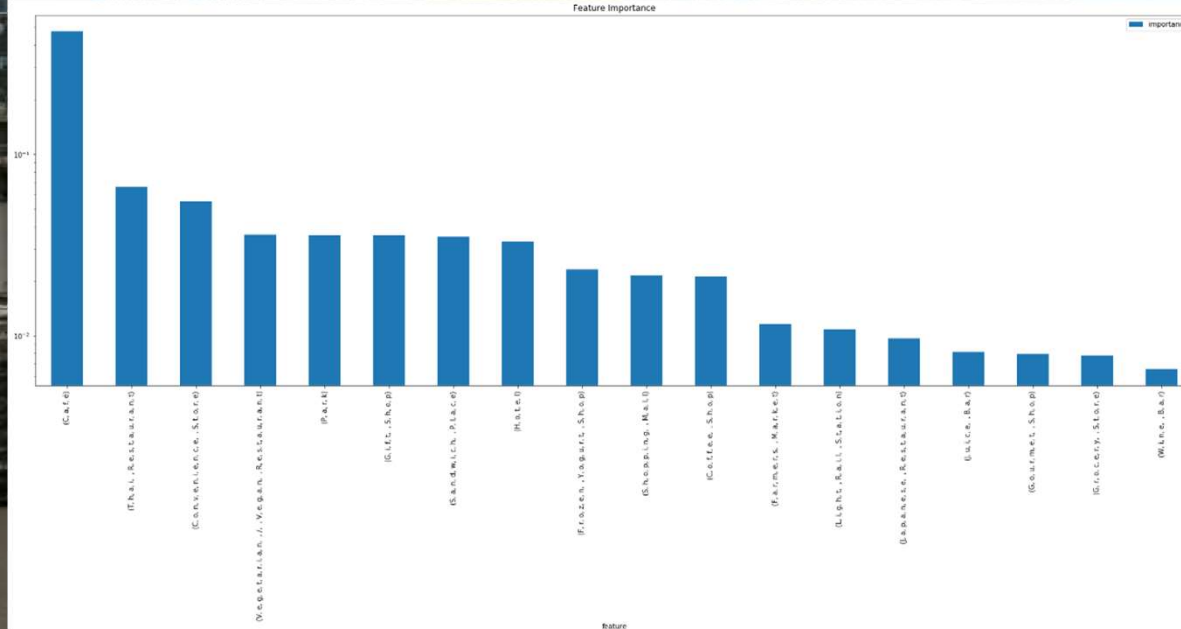
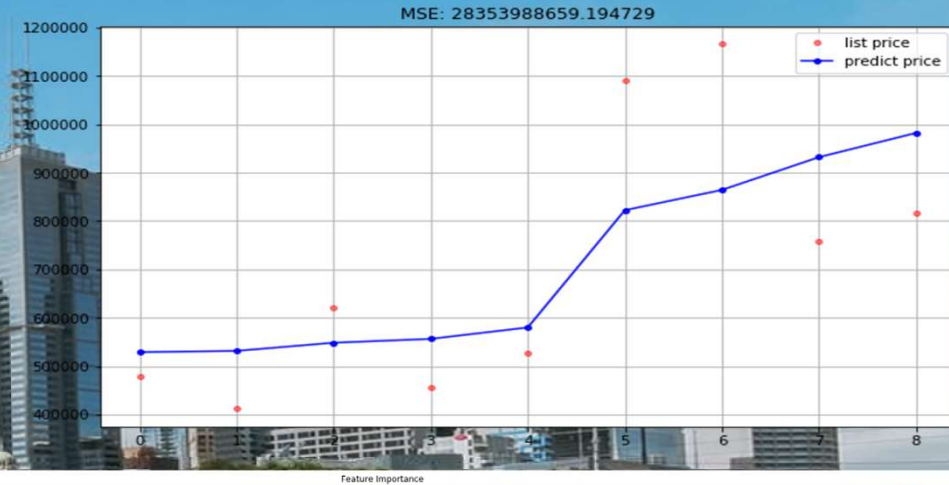
	Regression type	R <sup>2</sup>	MAE	MSE	RMS E
Southern Metropolitan Region	Linear	0.49	438260	4.058E+11	637025
		0	669042	7.945E+11	891349
		0.39	517633	4.871E+11	697946
		0.02	662206	7.787E+11	882458
	MLR	0.63	365962	2.924E+11	540745
	Polynomial Pipe	0.68	338543	2.509E+11	500889
	Ridge	0.69	353303	2.807E+11	529781
	Lasso	0.63	362870	2.893E+11	537895
	GBR	0.75	314712	2.317E+11	481312

From the table, it can be seen that Ridge & Polynomial Pipe regression outperform other regression models with an exception of using GBR, where principal component analysis was performed

# Random Forest Regression

Relationship between venue category and average property price was determined by Random Forest Regression resulting in ( $R^2$ ) of  $\sim 0.58$

Below is the plot showing venue with its importance with respect to average property price;





Results

	Regression type	R <sup>2</sup>	MAE	MSE	RMSE	Yhat		Regression type	R <sup>2</sup>	MAE	MSE	RMSE	Yhat		Regression type	R <sup>2</sup>	MAE	MSE	RMSE	Yhat		Regression type	R <sup>2</sup>	MAE	MSE	RMSE	Yhat																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
Melbourne	Linear	0.23	407809	3.474E+11	589435	724427	House type in all regions Melbourne	Linear	0.12	482321	4.4435E+11	666598	829940	Unit type in all regions in Melbourne	Linear	0.27	139612	4.43E+10	210393	601713	Townhouse type in all regions in Melbourne	Linear	0.19	244730	1.219E+11	349186	701027	2 rooms																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
		0.05	456927	4.264E+11	653013	882389			0.22	425552	3.525E+11	593727	1238825			0	479036	4.507E+11	671370	1081777			0.16	457295	4.2457E+11	651592	929938	0.173580	6E+10	245034	558879	0.02	279363	1.475E+11	384038	808019	20 km distance	0.17	468515	4.1831E+11	646768	1330695	0.22	154869	4.68E+10	216363	830029	0.2	245754	1.201E+11	346522	968423	2 bathroom	0.01	504354	4.9846E+11	706018	990582	0.01	172374	6E+10	244951	669846	0.01	279083	1.495E+11	386711	1029662	year 2019	0.45	359884	2.7484E+11	524248	517153	0.42	128069	3.5E+10	187039	666461	0.34	221245	9.857E+10	313953	641096		Polynomial Pipe	0.53	327849	2.3631E+11	486122	350900	Polynomial Pipe	0.48	123458	3.13E+10	177005	616668	Polynomial Pipe	0.4	209798	8.963E+10	299386	558819		Ridge	0.55	284587	1.859E+11	431125		Ridge	0.54	323986	2.4289E+11	492841		Ridge	0.53	123549	2.71E+10	164664		Ridge	0.39	223910	9.587E+10	309634			Lasso	0.52	308633	2.135E+11	462028		Lasso	0.46	358092	2.8476E+11	533628		Lasso	0.39	134973	3.66E+10	191342		Lasso	0.35	234859	1.019E+11	319204			GBR	0.62	270466	1.587E+11	398392		GBR	0.59	308234	2.19E+11	467972		GBR	0.49	126081	3.05E+10	174704		GBR	0.45	210631	8.81E+10	296813			Southern Metropolitan Region	Linear	0.49	438260	4.058E+11	637025	886874	House Type Southern Metropolitan Region	Linear	0.26	513922	5.378E+11	733212	1197267	Unit type in SM Region	Linear	0.36	141024	4.74E+10	217751	676644	Townhouse type in S M Region	Linear	0.18	283968	1.613E+11	401599	901739		0	669042	7.945E+11	891349	1365098	0.06	589097	6.7955E+11	824346	1274496	0	190323	7.45E+10	273035	661398	0.06	300037	1.852E+11	430300	925264		0.39	517633	4.871E+11	697946	1651077	0.3	505210	5.1025E+11	714320	1930511	0.25	167457	5.6E+10	236562	924835	0.16	287498	1.641E+11	405077	1200950		0.02	662206	7.787E+11	882458	1961055	0.01	610849	7.1993E+11	848487	2170160	0.01	189352	7.41E+10	272278	756801	0	310288	1.921E+11	438272	1456648		MLR	0.63	365962	2.924E+11	540743	522907	MLR	0.48	422645	3.7598E+11	613172	532558	MLR	0.49	129300	3.8E+10	194997	770772	MLR	0.37	238117	1.229E+11	350571	672380		Polynomial Pipe	0.68	338543	2.509E+11	500889	627889	Polynomial Pipe	0.61	371753	2.858E+11	534598	870663	Polynomial Pipe	0.53	123672	3.48E+10	186467	925145	Polynomial Pipe	0.48	217177	1.027E+11	320546	681997		Ridge	0.69	353303	2.807E+11	529781		Ridge	0.61	388576	3.0172E+11	549292		Ridge	0.64	128885	2.95E+10	171897		Ridge	0.44	248795	1.287E+11	358786			Lasso	0.63	362870	2.893E+11	537895		Lasso	0.48	386729	2.6819E+11	517867		Lasso	0.38	140583	4.92E+10	221771		Lasso	0.35	268272	1.507E+11	388232			GBR	0.75	314712	2.317E+11	481312		GBR	0.61	376565	3.0421E+11	551548		GBR	0.65	126782	2.86E+10	169070		GBR	0.53	239904	1.083E+11	329064			All types in Northern Metropolitan Region	Linear	0.13	307566	1.71E+11	413471	725713	House Type Northern Metropolitan Region	Linear	0.02	337502	2.0372E+11	451358	884373	Unit type in NM Region	Linear	0.22	124659	3.52E+10	187724	559261	Townhouse type in N M Region	Linear	0.04	192099	6.234E+10	249679	703287		0.16	280010	1.64E+11	405030	570957	0.07	318820	1.825E+11	427244	975305	0	321918	1.957E+11	442327	866567	0.39	239355	1.27E+11	356370	557449	0.08	139802	4.13E+10	203115	330682	0.49	132854	3.301E+10	181679	294654		0.03	334647	2.0248E+11	449980	1036375	0.24	126641	3.41E+10	184537	774839	0.12	183837	5.669E+10	238089	811421		0.02	329994	2.0491E+11	452669	795834	0.01	144254	4.47E+10	211448	611403	0.01	196135	6.427E+10	253522	819872		MLR	0.61	191768	7.717E+10	277798	331443	MLR	0.57	196862	8.5058E+10	291646	382061	MLR	0.51	100431	2.21E+10	148707	520936	MLR	0.66	109784	2.177E+10	147552	411856		Polynomial Pipe	0.69	167021	6E+10	244950	492378	Polynomial Pipe	0.7	167918	6.3257E+10	251509	400445	Polynomial Pipe	0.63	90004	1.68E+10	129569	425076	Polynomial Pipe	0.75	94928	1.602E+10	126574	559303		Ridge	0.7	151614	4.511E+10	212291		Ridge	0.65	182592	7.7665E+10	278683		Ridge	0.6	91011	1.95E+10	139656		Ridge	0.61	117792	2.646E+10	162679			Lasso	0.61	195388	7.266E+10	269551		Lasso	0.61	195348	7.4304E+10	272588		Lasso	0.53	94326	1.88E+10	137100		Lasso	0.58	115083	1.953E+10	139734			PCA	0.72	141530	4.181E+10	204465		GBR	0.7	167309	6.5795E+10	256504		GBR	0.56	92116	2.16E+10	146907		GBR	0.77	89951	1.253E+10	111941																								
		0.22	425552	3.525E+11	593727	1238825			0	479036	4.507E+11	671370	1081777			0.16	457295	4.2457E+11	651592	929938			0.173580	6E+10	245034	558879	0.02	279363	1.475E+11	384038	808019	20 km distance	0.17	468515	4.1831E+11	646768	1330695	0.22	154869	4.68E+10	216363	830029	0.2	245754	1.201E+11	346522	968423	2 bathroom	0.01	504354	4.9846E+11	706018	990582	0.01	172374	6E+10	244951	669846	0.01	279083	1.495E+11	386711	1029662	year 2019	0.45	359884	2.7484E+11	524248	517153	0.42	128069	3.5E+10	187039	666461	0.34	221245	9.857E+10	313953	641096		Polynomial Pipe	0.53	327849	2.3631E+11	486122	350900	Polynomial Pipe	0.48	123458	3.13E+10	177005	616668	Polynomial Pipe	0.4	209798	8.963E+10	299386	558819		Ridge	0.55	284587	1.859E+11	431125		Ridge	0.54	323986	2.4289E+11	492841		Ridge	0.53	123549	2.71E+10	164664		Ridge	0.39	223910	9.587E+10	309634			Lasso	0.52	308633	2.135E+11	462028		Lasso	0.46	358092	2.8476E+11	533628		Lasso	0.39	134973	3.66E+10	191342		Lasso	0.35	234859	1.019E+11	319204			GBR	0.62	270466	1.587E+11	398392		GBR	0.59	308234	2.19E+11	467972		GBR	0.49	126081	3.05E+10	174704		GBR	0.45	210631	8.81E+10	296813			Southern Metropolitan Region	Linear	0.49	438260	4.058E+11			637025	886874	House Type Southern Metropolitan Region	Linear	0.26			513922	5.378E+11	733212	1197267	Unit type in SM Region			Linear	0.36	141024	4.74E+10	217751			676644	Townhouse type in S M Region	Linear	0.18	283968	1.613E+11	401599	901739		0	669042	7.945E+11	891349	1365098	0.06	589097	6.7955E+11	824346	1274496	0	190323	7.45E+10	273035	661398	0.06	300037	1.852E+11	430300	925264		0.39	517633	4.871E+11	697946	1651077	0.3	505210	5.1025E+11	714320	1930511	0.25	167457	5.6E+10	236562	924835	0.16	287498	1.641E+11	405077	1200950		0.02	662206	7.787E+11	882458	1961055	0.01	610849	7.1993E+11	848487	2170160	0.01	189352	7.41E+10	272278	756801	0	310288	1.921E+11	438272	1456648		MLR	0.63	365962	2.924E+11	540743	522907	MLR	0.48	422645	3.7598E+11	613172	532558	MLR	0.49	129300	3.8E+10	194997	770772	MLR	0.37	238117	1.229E+11	350571	672380		Polynomial Pipe	0.68	338543	2.509E+11	500889	627889	Polynomial Pipe	0.61	371753	2.858E+11	534598	870663	Polynomial Pipe	0.53	123672	3.48E+10	186467	925145	Polynomial Pipe	0.48	217177	1.027E+11	320546	681997		Ridge	0.69	353303	2.807E+11	529781		Ridge	0.61	388576	3.0172E+11	549292		Ridge	0.64	128885	2.95E+10	171897		Ridge	0.44	248795	1.287E+11	358786			Lasso	0.63	362870	2.893E+11	537895		Lasso	0.48	386729	2.6819E+11	517867		Lasso	0.38	140583	4.92E+10	221771		Lasso	0.35	268272	1.507E+11	388232			GBR	0.75	314712	2.317E+11	481312		GBR	0.61	376565	3.0421E+11	551548		GBR	0.65	126782	2.86E+10	169070		GBR	0.53	239904	1.083E+11			329064			All types in Northern Metropolitan Region	Linear			0.13	307566	1.71E+11	413471	725713			House Type Northern Metropolitan Region	Linear	0.02	337502	2.0372E+11			451358	884373	Unit type in NM Region	Linear	0.22	124659	3.52E+10	187724	559261	Townhouse type in N M Region	Linear	0.04	192099	6.234E+10	249679	703287		0.16	280010	1.64E+11	405030	570957	0.07	318820	1.825E+11	427244	975305	0	321918	1.957E+11	442327	866567	0.39	239355	1.27E+11	356370	557449	0.08	139802	4.13E+10	203115	330682	0.49	132854	3.301E+10	181679	294654		0.03	334647	2.0248E+11	449980	1036375	0.24	126641	3.41E+10	184537	774839	0.12	183837	5.669E+10	238089	811421		0.02	329994	2.0491E+11	452669	795834	0.01	144254	4.47E+10	211448	611403	0.01	196135	6.427E+10	253522	819872		MLR	0.61	191768	7.717E+10	277798	331443	MLR	0.57	196862	8.5058E+10	291646	382061	MLR	0.51	100431	2.21E+10	148707	520936	MLR	0.66	109784	2.177E+10	147552	411856		Polynomial Pipe	0.69	167021	6E+10	244950	492378	Polynomial Pipe	0.7	167918	6.3257E+10	251509	400445	Polynomial Pipe	0.63	90004	1.68E+10	129569	425076	Polynomial Pipe	0.75	94928	1.602E+10	126574	559303		Ridge	0.7	151614	4.511E+10	212291		Ridge	0.65	182592	7.7665E+10	278683		Ridge	0.6	91011	1.95E+10	139656		Ridge	0.61	117792	2.646E+10	162679			Lasso	0.61	195388	7.266E+10	269551		Lasso	0.61	195348	7.4304E+10	272588		Lasso	0.53	94326	1.88E+10	137100		Lasso	0.58	115083	1.953E+10	139734			PCA	0.72	141530	4.181E+10	204465		GBR	0.7	167309	6.5795E+10	256504		GBR	0.56	92116	2.16E+10	146907		GBR	0.77	89951	1.253E+10	111941													
		0	479036	4.507E+11	671370	1081777			0.16	457295	4.2457E+11	651592	929938			0.173580	6E+10	245034	558879	0.02			279363	1.475E+11	384038	808019	20 km distance	0.17	468515	4.1831E+11	646768	1330695	0.22	154869	4.68E+10	216363	830029	0.2	245754	1.201E+11	346522	968423	2 bathroom	0.01	504354	4.9846E+11	706018	990582	0.01	172374	6E+10	244951	669846	0.01	279083	1.495E+11	386711	1029662	year 2019	0.45	359884	2.7484E+11	524248	517153	0.42	128069	3.5E+10	187039	666461	0.34	221245	9.857E+10	313953	641096		Polynomial Pipe	0.53	327849	2.3631E+11	486122	350900	Polynomial Pipe	0.48	123458	3.13E+10	177005	616668	Polynomial Pipe	0.4	209798	8.963E+10	299386	558819		Ridge	0.55	284587	1.859E+11	431125		Ridge	0.54	323986	2.4289E+11	492841		Ridge	0.53	123549	2.71E+10	164664		Ridge	0.39	223910	9.587E+10	309634			Lasso	0.52	308633	2.135E+11	462028		Lasso	0.46	358092	2.8476E+11	533628		Lasso	0.39	134973	3.66E+10	191342		Lasso	0.35	234859	1.019E+11	319204			GBR	0.62	270466	1.587E+11	398392		GBR	0.59	308234	2.19E+11	467972		GBR	0.49	126081	3.05E+10	174704		GBR	0.45	210631	8.81E+10	296813			Southern Metropolitan Region	Linear	0.49	438260	4.058E+11			637025	886874	House Type Southern Metropolitan Region			Linear	0.26			513922			5.378E+11	733212	1197267	Unit type in SM Region					Linear	0.36	141024	4.74E+10			217751			676644	Townhouse type in S M Region	Linear	0.18	283968	1.613E+11	401599	901739		0	669042	7.945E+11	891349	1365098	0.06	589097	6.7955E+11	824346	1274496	0	190323	7.45E+10	273035	661398	0.06	300037	1.852E+11	430300	925264		0.39	517633	4.871E+11	697946	1651077	0.3	505210	5.1025E+11	714320	1930511	0.25	167457	5.6E+10	236562	924835	0.16	287498	1.641E+11	405077	1200950		0.02	662206	7.787E+11	882458	1961055	0.01	610849	7.1993E+11	848487	2170160	0.01	189352	7.41E+10	272278	756801	0	310288	1.921E+11	438272	1456648		MLR	0.63	365962	2.924E+11	540743	522907	MLR	0.48	422645	3.7598E+11	613172	532558	MLR	0.49	129300	3.8E+10	194997	770772	MLR	0.37	238117	1.229E+11	350571	672380		Polynomial Pipe	0.68	338543	2.509E+11	500889	627889	Polynomial Pipe	0.61	371753	2.858E+11	534598	870663	Polynomial Pipe	0.53	123672	3.48E+10	186467	925145	Polynomial Pipe	0.48	217177	1.027E+11	320546	681997		Ridge	0.69	353303	2.807E+11	529781		Ridge	0.61	388576	3.0172E+11	549292		Ridge	0.64	128885	2.95E+10	171897		Ridge	0.44	248795	1.287E+11	358786			Lasso	0.63	362870	2.893E+11	537895		Lasso	0.48	386729	2.6819E+11	517867		Lasso	0.38	140583	4.92E+10	221771		Lasso	0.35	268272	1.507E+11	388232			GBR	0.75	314712	2.317E+11	481312		GBR	0.61	376565	3.0421E+11	551548		GBR	0.65	126782	2.86E+10	169070		GBR			0.53	239904	1.083E+11					329064			All types in Northern Metropolitan Region	Linear					0.13	307566	1.71E+11			413471	725713			House Type Northern Metropolitan Region	Linear	0.02	337502	2.0372E+11			451358	884373	Unit type in NM Region	Linear	0.22	124659	3.52E+10	187724	559261	Townhouse type in N M Region	Linear	0.04	192099	6.234E+10	249679	703287		0.16	280010	1.64E+11	405030	570957	0.07	318820	1.825E+11	427244	975305	0	321918	1.957E+11	442327	866567	0.39	239355	1.27E+11	356370	557449	0.08	139802	4.13E+10	203115	330682	0.49	132854	3.301E+10	181679	294654		0.03	334647	2.0248E+11	449980	1036375	0.24	126641	3.41E+10	184537	774839	0.12	183837	5.669E+10	238089	811421		0.02	329994	2.0491E+11	452669	795834	0.01	144254	4.47E+10	211448	611403	0.01	196135	6.427E+10	253522	819872		MLR	0.61	191768	7.717E+10	277798	331443	MLR	0.57	196862	8.5058E+10	291646	382061	MLR	0.51	100431	2.21E+10	148707	520936	MLR	0.66	109784	2.177E+10	147552	411856		Polynomial Pipe	0.69	167021	6E+10	244950	492378	Polynomial Pipe	0.7	167918	6.3257E+10	251509	400445	Polynomial Pipe	0.63	90004	1.68E+10	129569	425076	Polynomial Pipe	0.75	94928	1.602E+10	126574	559303		Ridge	0.7	151614	4.511E+10	212291		Ridge	0.65	182592	7.7665E+10	278683		Ridge	0.6	91011	1.95E+10	139656		Ridge	0.61	117792	2.646E+10	162679			Lasso	0.61	195388	7.266E+10	269551		Lasso	0.61	195348	7.4304E+10	272588		Lasso	0.53	94326	1.88E+10	137100		Lasso	0.58	115083	1.953E+10	139734			PCA	0.72	141530	4.181E+10	204465		GBR	0.7	167309	6.5795E+10	256504		GBR	0.56	92116	2.16E+10	146907		GBR	0.77	89951	1.253E+10	111941		
		0.16	457295	4.2457E+11	651592	929938			0.173580	6E+10	245034	558879	0.02			279363	1.475E+11	384038	808019	20 km distance																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
	0.17	468515	4.1831E+11	646768	1330695	0.22		154869	4.68E+10	216363	830029	0.2	245754		1.201E+11	346522	968423	2 bathroom																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
	0.01	504354	4.9846E+11	706018	990582	0.01		172374	6E+10	244951	669846	0.01	279083		1.495E+11	386711	1029662	year 2019																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
	0.45	359884	2.7484E+11	524248	517153	0.42		128069	3.5E+10	187039	666461	0.34	221245		9.857E+10	313953	641096																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
	Polynomial Pipe	0.53	327849	2.3631E+11	486122	350900		Polynomial Pipe	0.48	123458	3.13E+10	177005	616668		Polynomial Pipe	0.4	209798	8.963E+10	299386	558819																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
Ridge	0.55	284587	1.859E+11	431125		Ridge	0.54	323986	2.4289E+11	492841		Ridge	0.53	123549	2.71E+10	164664		Ridge	0.39	223910	9.587E+10	309634																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
Lasso	0.52	308633	2.135E+11	462028		Lasso	0.46	358092	2.8476E+11	533628		Lasso	0.39	134973	3.66E+10	191342		Lasso	0.35	234859	1.019E+11	319204																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
GBR	0.62	270466	1.587E+11	398392		GBR	0.59	308234	2.19E+11	467972		GBR	0.49	126081	3.05E+10	174704		GBR	0.45	210631	8.81E+10	296813																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
Southern Metropolitan Region	Linear	0.49	438260	4.058E+11	637025	886874	House Type Southern Metropolitan Region	Linear	0.26	513922	5.378E+11	733212	1197267	Unit type in SM Region	Linear	0.36	141024	4.74E+10	217751	676644	Townhouse type in S M Region	Linear	0.18	283968	1.613E+11	401599	901739																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
		0	669042	7.945E+11	891349	1365098			0.06	589097	6.7955E+11	824346	1274496			0	190323	7.45E+10	273035	661398			0.06	300037	1.852E+11	430300	925264																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
		0.39	517633	4.871E+11	697946	1651077			0.3	505210	5.1025E+11	714320	1930511			0.25	167457	5.6E+10	236562	924835			0.16	287498	1.641E+11	405077	1200950																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
		0.02	662206	7.787E+11	882458	1961055			0.01	610849	7.1993E+11	848487	2170160			0.01	189352	7.41E+10	272278	756801			0	310288	1.921E+11	438272	1456648																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
	MLR	0.63	365962	2.924E+11	540743	522907		MLR	0.48	422645	3.7598E+11	613172	532558		MLR	0.49	129300	3.8E+10	194997	770772		MLR	0.37	238117	1.229E+11	350571	672380																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
	Polynomial Pipe	0.68	338543	2.509E+11	500889	627889		Polynomial Pipe	0.61	371753	2.858E+11	534598	870663		Polynomial Pipe	0.53	123672	3.48E+10	186467	925145		Polynomial Pipe	0.48	217177	1.027E+11	320546	681997																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
	Ridge	0.69	353303	2.807E+11	529781			Ridge	0.61	388576	3.0172E+11	549292			Ridge	0.64	128885	2.95E+10	171897			Ridge	0.44	248795	1.287E+11	358786																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
	Lasso	0.63	362870	2.893E+11	537895			Lasso	0.48	386729	2.6819E+11	517867			Lasso	0.38	140583	4.92E+10	221771			Lasso	0.35	268272	1.507E+11	388232																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
	GBR	0.75	314712	2.317E+11	481312			GBR	0.61	376565	3.0421E+11	551548			GBR	0.65	126782	2.86E+10	169070			GBR	0.53	239904	1.083E+11	329064																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
All types in Northern Metropolitan Region	Linear	0.13	307566	1.71E+11	413471	725713	House Type Northern Metropolitan Region	Linear	0.02	337502	2.0372E+11	451358	884373	Unit type in NM Region	Linear	0.22	124659	3.52E+10	187724	559261	Townhouse type in N M Region	Linear	0.04	192099	6.234E+10	249679	703287																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
		0.16	280010	1.64E+11	405030	570957			0.07	318820	1.825E+11	427244	975305			0	321918	1.957E+11	442327	866567			0.39	239355	1.27E+11	356370	557449	0.08	139802	4.13E+10	203115	330682	0.49	132854	3.301E+10	181679	294654		0.03	334647	2.0248E+11	449980	1036375	0.24	126641	3.41E+10	184537	774839	0.12	183837	5.669E+10	238089	811421		0.02	329994	2.0491E+11	452669	795834	0.01	144254	4.47E+10	211448	611403	0.01	196135	6.427E+10	253522	819872		MLR	0.61	191768	7.717E+10	277798	331443	MLR	0.57	196862	8.5058E+10	291646	382061	MLR	0.51	100431	2.21E+10	148707	520936	MLR	0.66	109784	2.177E+10	147552	411856		Polynomial Pipe	0.69	167021	6E+10	244950	492378	Polynomial Pipe	0.7	167918	6.3257E+10	251509	400445	Polynomial Pipe	0.63	90004	1.68E+10	129569	425076	Polynomial Pipe	0.75	94928	1.602E+10	126574	559303		Ridge	0.7	151614	4.511E+10	212291		Ridge	0.65	182592	7.7665E+10	278683		Ridge	0.6	91011	1.95E+10	139656		Ridge	0.61	117792	2.646E+10	162679			Lasso	0.61	195388	7.266E+10	269551		Lasso	0.61	195348	7.4304E+10	272588		Lasso	0.53	94326	1.88E+10	137100		Lasso	0.58	115083	1.953E+10	139734			PCA	0.72	141530	4.181E+10	204465		GBR	0.7	167309	6.5795E+10	256504		GBR	0.56	92116	2.16E+10	146907		GBR	0.77	89951	1.253E+10	111941																																																																																																																																																																																																																																																																																																																																																																																																																																																										
		0.07	318820	1.825E+11	427244	975305			0	321918	1.957E+11	442327	866567			0.39	239355	1.27E+11	356370	557449			0.08	139802	4.13E+10	203115	330682	0.49	132854	3.301E+10	181679	294654		0.03	334647	2.0248E+11	449980	1036375	0.24	126641	3.41E+10	184537	774839	0.12	183837	5.669E+10	238089	811421		0.02	329994	2.0491E+11	452669	795834	0.01	144254	4.47E+10	211448	611403	0.01	196135	6.427E+10	253522	819872		MLR	0.61	191768	7.717E+10	277798	331443	MLR	0.57	196862	8.5058E+10	291646	382061	MLR	0.51	100431	2.21E+10	148707	520936	MLR	0.66	109784	2.177E+10	147552	411856		Polynomial Pipe	0.69	167021	6E+10	244950	492378	Polynomial Pipe	0.7	167918	6.3257E+10	251509	400445	Polynomial Pipe	0.63	90004	1.68E+10	129569	425076	Polynomial Pipe	0.75	94928	1.602E+10	126574	559303		Ridge	0.7	151614	4.511E+10	212291		Ridge	0.65	182592	7.7665E+10	278683		Ridge	0.6	91011	1.95E+10	139656		Ridge	0.61	117792	2.646E+10	162679			Lasso	0.61	195388	7.266E+10	269551		Lasso	0.61	195348	7.4304E+10	272588		Lasso	0.53	94326	1.88E+10	137100		Lasso	0.58	115083	1.953E+10	139734			PCA	0.72	141530	4.181E+10	204465		GBR	0.7	167309	6.5795E+10	256504		GBR	0.56	92116	2.16E+10	146907		GBR	0.77	89951	1.253E+10	111941																																																																																																																																																																																																																																																																																																																																																																																																																																																															
		0	321918	1.957E+11	442327	866567			0.39	239355	1.27E+11	356370	557449			0.08	139802	4.13E+10	203115	330682			0.49	132854	3.301E+10	181679	294654																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
	0.03	334647	2.0248E+11	449980	1036375	0.24		126641	3.41E+10	184537	774839	0.12	183837		5.669E+10	238089	811421																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
	0.02	329994	2.0491E+11	452669	795834	0.01		144254	4.47E+10	211448	611403	0.01	196135		6.427E+10	253522	819872																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
	MLR	0.61	191768	7.717E+10	277798	331443		MLR	0.57	196862	8.5058E+10	291646	382061		MLR	0.51	100431	2.21E+10	148707	520936		MLR	0.66	109784	2.177E+10	147552	411856																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
	Polynomial Pipe	0.69	167021	6E+10	244950	492378		Polynomial Pipe	0.7	167918	6.3257E+10	251509	400445		Polynomial Pipe	0.63	90004	1.68E+10	129569	425076		Polynomial Pipe	0.75	94928	1.602E+10	126574	559303																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
	Ridge	0.7	151614	4.511E+10	212291			Ridge	0.65	182592	7.7665E+10	278683			Ridge	0.6	91011	1.95E+10	139656			Ridge	0.61	117792	2.646E+10	162679																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
Lasso	0.61	195388	7.266E+10	269551		Lasso	0.61	195348	7.4304E+10	272588		Lasso	0.53	94326	1.88E+10	137100		Lasso	0.58	115083	1.953E+10	139734																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
PCA	0.72	141530	4.181E+10	204465		GBR	0.7	167309	6.5795E+10	256504		GBR	0.56	92116	2.16E+10	146907		GBR	0.77	89951	1.253E+10	111941																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					

## Results

- Histogram for all property type was generated in order to provide an insight of the pricing frequency in the city of Melbourne in general which lies between \$ 800,000 - 120,000,0.
- Box plot was generated to identify the relationship between price & categorical variables but there seems an overlap between all variables, hence these variables can't be good predictors of price.
- Heat map & Pair plot suggest the strong to moderate relationship of price with number of rooms, distance from CBD, number of bathrooms & year built.
- Line plot between mean price of all three types of properties in Melbourne suggests that buyers should be interested in buying House in coming years due to dramatic change (~ \$100,000/year) in price.
- Line plot between mean price of all 3 types of properties in Southern Metropolitan region suggests that sellers should be interested in selling Houses in coming years due to dramatic change (increased by ~ \$100,000/year) in price.
- Line plot between mean price of all 3 types of properties in Northern Metropolitan region suggests that sellers should be interested in selling Townhouses in coming years due to dramatic change in price.
- GBR gives maximum  $R^2$  of 0.62, 0.59 & 0.45 when applied on all variables for all type of properties, for type = house & for type = townhouse in Melbourne, respectively.
- Ridge regression gives maximum  $R^2$  of 0.53 when applied on all variables for type = unit in Melbourne.
- GBR gives maximum  $R^2$  of 0.75, 0.61, 0.65 & 0.53 when applied on all variables for all type of properties, for type = house, for type = unit & for type = townhouse in Southern Metropolitan region.
- GBR gives maximum  $R^2$  of 0.72, 0.7 & 0.77 when applied on all variables for all type of properties, for type = house & for type = townhouse in Northern Metropolitan region.
- Maximum number of venues & unique venues category are concentrated in Northern Metropolitan region, where house & units are located.



## Discussion & Recommendation

- For any newcomer, the main area of interest for finding accommodation should be the neighborhood located in Northern Metropolitan region.
- Secondly, property type condo should be the focus due to its affordability in comparison to other property types.
- Thirdly, number of bathrooms requirement to be optimized as per family need as it has a moderate to strong association with property price.
- Lastly, consider those neighborhoods which are without or at least have minimum number of venues such as cafes, dessert shops & fast food restaurants as they have a major impact on property price.





## Conclusion

- Based on data visualization, neighborhoods closer to the Southern Metropolitan region in Melbourne are the expensive ones possibly due to less venue types.
- Based on data visualization, neighborhoods closer to the Northern Metropolitan region in Melbourne are the least expensive ones possibly due to greater number of venue types.
- With respect to number of venue types and category, price for townhouse & house in Northern Metropolitan region shows reasonably good relationship with  $R^2$  value of 0.58 & 0.4, respectively.
- With respect to modelling including Multilinear, Polynomial, Lasso and Ridge regression shows similar results in terms of regression coefficient and mean squared error, with exceptions of Unit & Townhouse types in Southern Metropolitan region.
- Since regression techniques including linear and polynomial regression didn't show any association of venue category data with neighborhood average property price. In this regard, random forest regression with feature importance was performed which improved regression coefficient and showed feature importance with respect to average property price of neighborhoods.
- Based on feature importance results; cafes, Australian & Japanese Restaurants, light rail station and convenient stores are considered as top three venues that have major impact on average property price of property in Melbourne.