

# Bangkok Housing Price Prediction

For buyers and sellers

# Problems

Do you ever have a problem when moving to a new home ?

Housing price in Bangkok is normally fluctuated due to many factors, some factors are considered to be positive for the price, but is that always true ?

The data:

# Bangkok Housing Price

Column	Type	Description
id	int	ID of selling item
province	string	province name: this dataset only includes Bangkok,Samut Prakan and Nonthaburi
district	string	district name
subdistrict	string	subdtistrict name
address	string	address e.g. street name, area name, soi number
property_type	string	type of the house: Condo, Townhouse or Detached House
total_units	float	the number of rooms/houses that the condo/village has
bedrooms	int	the number of bedrooms
baths	int	the number of baths
floor_area	float	total area of inside floor [m <sup>2</sup> ]
floor_level	int	floor level of the room
land_area	float	total area of the land [m <sup>2</sup> ]
latitude	float	latitude of the house
longitude	float	longitude of the house
nearby_stations	int	the number of nearby stations (within 1km)
nearby_station_distance	list	list of (station name, distance[m]). Each station name consists of station ID, station name, and Line such as "E4 Asok BTS"
nearby_bus_stops	int	the number of nearby bus stops
nearby_supermarkets	int	the number of nearby supermarkets
nearby_shops	int	the number of nearby shops
year_built	int	year built
month_built	string	month built: January-December
facilities	list	list of facilities
price	float	selling price

# Considering about the information ...

## Useless

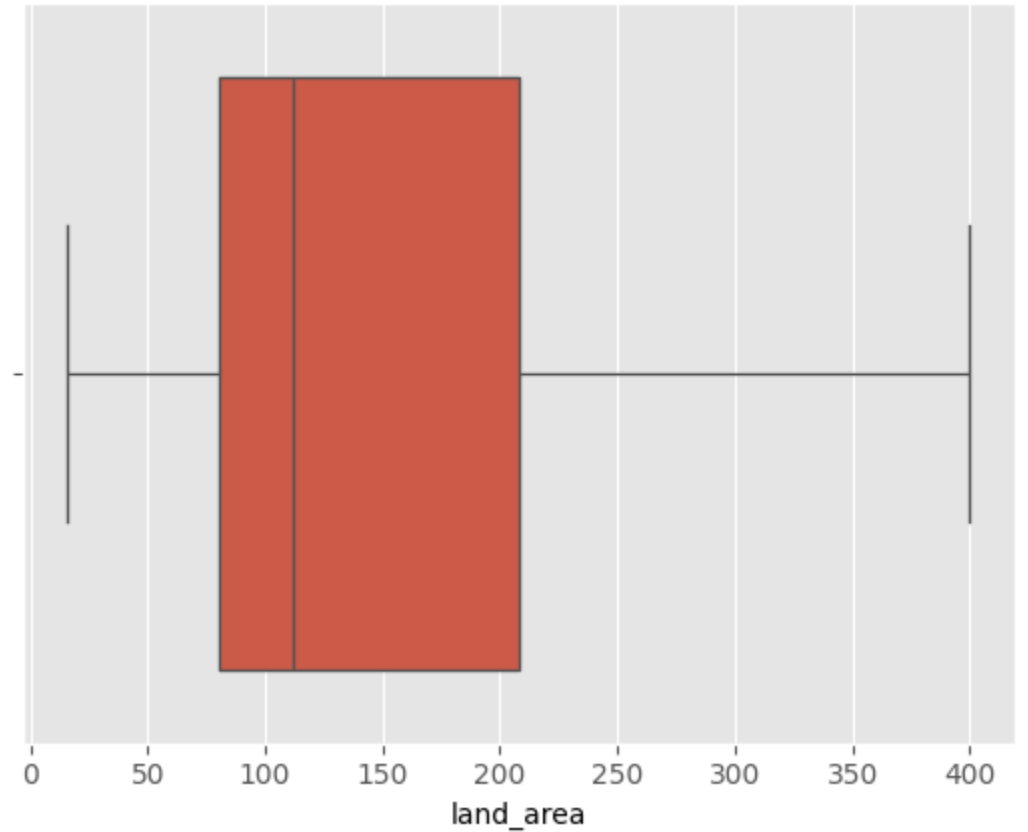
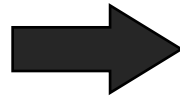
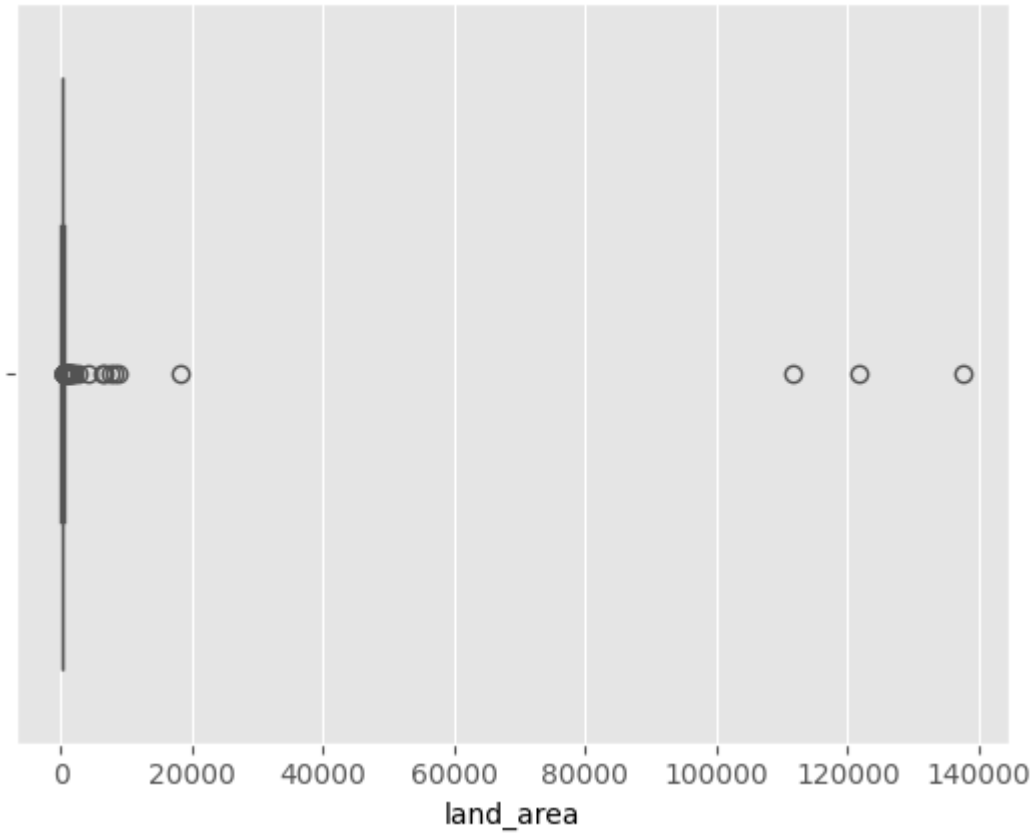
- Id
- Month built
- Address

## Redundant

- Province
- Latitude
- Longitude

Column	Type	Description
district	string	district name
subdistrict	string	subdtistrict name
property_type	string	type of the house: Condo, Townhouse or Detached House
total_units	float	the number of rooms/houses that the condo/village has
bedrooms	int	the number of bedrooms
baths	int	the number of baths
floor_area	float	total area of inside floor [m <sup>2</sup> ]
floor_level	int	floor level of the room
land_area	float	total area of the land [m <sup>2</sup> ]
nearby_stations	int	the number of nearby stations (within 1km)
nearby_station_distance	list	list of (station name, distance[m]). Each station name consists of station ID, station name, and Line such as "E4 Asok BTS"
nearby_bus_stops	int	the number of nearby bus stops
nearby_supermarkets	int	the number of nearby supermarkets
nearby_shops	int	the number of nearby shops
year_built	int	year built
facilities	list	list of facilities

# Outliers - land\_area (cutoff at 400 m<sup>2</sup>)

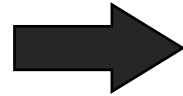




# For nearby\_station\_distance

- Custom transformation + normalization

	nearby_station_distance
0	[[E7 Ekkamai BTS, 270], [E6 Thong Lo BTS, 800]]
1	[[BL22 Sukhumvit MRT, 720], [BL21 Phetchaburi ...
2	[[E5 Phrom Phong BTS, 650], [BL23 Queen Siriki...
3	None
4	[[PP09 Yaek Nonthaburi 1 MRT, 10]]



	E7 Ekkamai BTS	E6 Thong Lo BTS	BL22 Sukhumvit MRT	BL21 Phetchaburi MRT
0	-1.118517	0.890493	-2.141975	-2.141975
1	-2.141975	-2.141975	0.587246	1.080022
2	-2.141975	-2.141975	-2.141975	-2.141975
3	-2.141975	-2.141975	-2.141975	-2.141975
4	-2.141975	-2.141975	-2.141975	-2.141975

# Data transformation summary

Column	Description	Transformation
floor_area	total area of inside floor [m <sup>2</sup> ]	log + polynomial degree 2
floor_level	floor level of the room	log + polynomial degree 2
land_area	total area of the land [m <sup>2</sup> ]	log + polynomial degree 2
total_units	the number of rooms/houses that the condo/village has	Polynomial degree 2
bedrooms	the number of bedrooms	Polynomial degree 2
baths	the number of baths	Polynomial degree 2
nearby_stations	the number of nearby stations (within 1km)	polynomial degree 2
nearby_supermarkets	the number of nearby supermarkets	polynomial degree 2
nearby_shops	the number of nearby shops	polynomial degree 2
year_built	year built	polynomial degree 2
facilities	list of facilities	Converted to facilities count - polynomial degree 2
nearby_bus_stops	the number of nearby bus stops	no transformation
nearby_station_distance	list of (station name, distance[m]). Each station name consists of station ID, station name, and Line such as "E4 Asok BTS"	Dummified-like (to normalized distance)
dist_subdist	district + subdistrict name concatenated	Dummified
property_type	type of the house: Condo, Townhouse or Detached House	Dummified



# Result

- Prediction model that
  - Can explain **79.4%** of data
  - The average error of **945K** baht approximately

---

Submission and Description

Public Score ⓘ



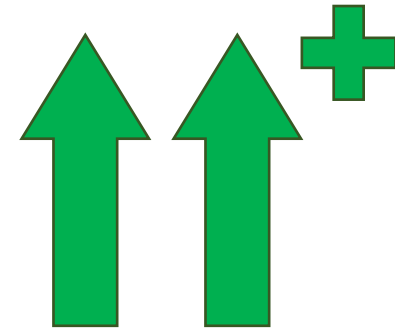
**submission.csv**

Complete · 6h ago · linear

**970438**

# Other findings

- The following factors are significantly affect housing price positively.
  - Floor area
  - Number of nearby stations (within 1km)
  - Number of bathrooms
  - Total area of the land
  - Year built
  - Number of bedrooms
- Locating neaby these stations will significantly raise the price as well.
  - BL11 Bang Sue MRT
  - BL32 Itsaraphap MRT
  - B9 Rama IX Bridge BRT



# Other findings

- Top 10 district/subdistrict that will have higher housing price according to the model.

District	Subdistrict
Bang Rak	Maha Phruettharam
Bang Rak	Suriyawong
Watthana	Khlong Tan Nuea
Pathum Wan	Rong Mueang
Watthana	Khlong Toei Nuea
Bang Rak	Si Phraya
Bang Kho Laem	Bang Kho Laem
Pathum Wan	Lumphini
Khlong Toei	Khlong Tan
Bang Rak	Silom