

# **CAPSTONE PROJECT MICROSOFT DATA & ARTIFICIAL INTELLIGENCE**

**FERZA REYALDI - DAI 003**

# MARIBISNIS

MariBisnis wants to know the business trend of selling homes in the United States.

The goal is to predict the price of a house and map out the distribution of the data.

# STEP-BY-STEP

01

**DATA UNDERSTANDING**

02

**CREATE WORKSPACE & RESOURCE**

03

**CREATE & DEPLOY PIPELINE**

04

**CREATE A REPORT IN POWER BI**

# DATA UNDERSTANDING

This dataset contains house sale prices in United States between May 2014 and May 2015.

# DATA UNDERSTANDING

Column name	Description	Column name	Description
<b>id</b>	a notation for a house	<b>date</b>	date house was sold
<b>price</b>	price is prediction target	<b>bedrooms</b>	number of bedrooms per house
<b>bathrooms</b>	number of bathrooms per bedrooms	<b>sqft_living</b>	square footage of the home
<b>sqft_lot</b>	square footage of the lot	<b>floors</b>	total floors (levels) in house
<b>waterfront</b>	house which has a view to waterfornt	<b>view</b>	has been viewed
<b>condition</b>	how good the condition is overall	<b>grade</b>	overall grade given to the housing unit, based on the grading system
<b>sqft_above</b>	square footage of house apart from basement	<b>sqft_basement</b>	square footage of the basemenet
<b>yr_built</b>	built year	<b>yr_renovated</b>	year when house was renovated
<b>zipcode</b>	zipcode	<b>lat</b>	latitude coordinate
<b>long</b>	longitude coordinate	<b>sqft_living15</b>	living room area in 2015 (implies-- some renovation), might or might not have affected the lotsize
<b>sqft_lot15</b>	lot size area in 2015 (implies-- some renovations)		

## CREATE WORKSPACE & RESOURCE

### Basics

Subscription	Azure for Students
Resource group	(New) ferza-ml
Region	Southeast Asia
Workspace name	capstone-ferza-dai
Storage account	(new) capstoneferzad7239751030
Key vault	(new) capstoneferzad6906677918
Application insights	(new) capstoneferzad2259061576
Container registry	None

Create Azure Machine Learning Workspace

# CREATE WORKSPACE & RESOURCE

Attributes	
Compute name	capstone-dai-ferza-ci
Compute type	Compute instance
Subscription ID	992f7a90-66f9-4c73-aa35-6deaa6bb6f0c
Resource group	ferza-ml
Workspace	capstone-ferza-dai
Region	southeastasia
Created by	Ferza Reyaldi

Create compute resources: compute instance

# CREATE WORKSPACE & RESOURCE

Attributes	Resource properties
<b>Compute name</b> capstonedairr-cc	<b>Virtual machine size</b> Standard_DS11_v2 (2 cores, 14 GB RAM, 28 GB disk)
<b>Resource ID</b> --	<b>Processing unit</b> <a href="#">CPU - Memory optimized</a>
<b>Compute type</b> Machine Learning compute	<b>OS Type</b> Linux
<b>Subscription ID</b> 992f7a90-66f9-4c73-aa35-6deaa6bb6f0c	<b>Virtual machine priority</b> Dedicated
<b>Resource group</b> ferza-ml	<b>Minimum number of nodes</b> 0
<b>Workspace</b> capstone-ferza-dai	<b>Maximum number of nodes</b> 2
<b>Region</b> southeastasia	<b>Idle seconds before scale down</b> 120
	<b>Virtual network/subnet</b> --

Create compute resources: compute cluster



## CREATE AND DEPLOY PIPELINE

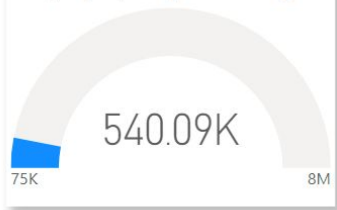
- Training pipeline
- Auto Inference Pipeline

## EVALUATE MODEL

Mean Absolute Error (MAE)	73445.9054
Root Mean Squared Error (RMSE)	122700.0282
Relative Squared Error (RSE)	0.1181
Relative Absolute Error (RAE)	0.3170
Coefficient of Determination	0.8819

# CREATE A REPORT IN POWER BI

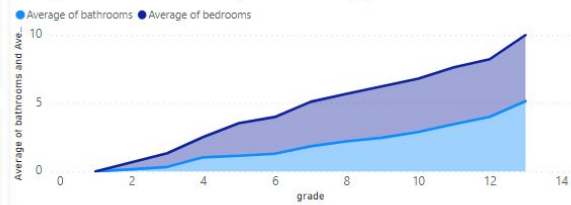
Average of price, Min of price and Max of price



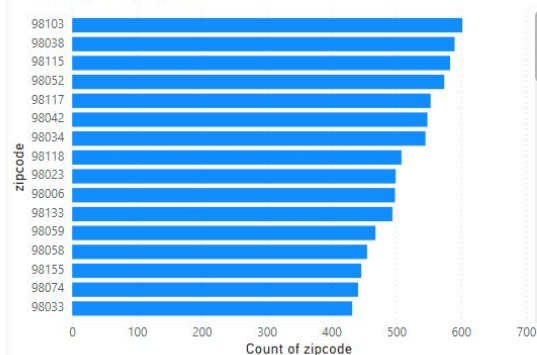
0  
Min of bathrooms  
8  
Max of bathrooms

0  
Min of bedrooms  
33  
Max of bedrooms

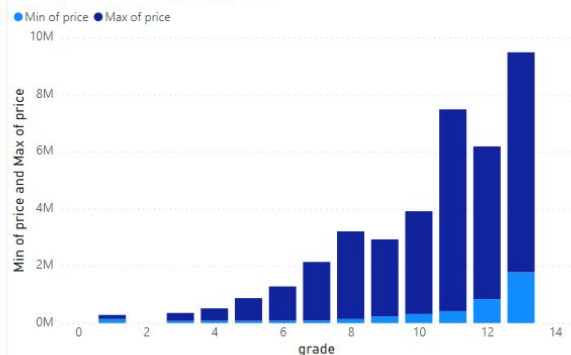
Average of bathrooms and Average of bedrooms by grade



Count of zipcode by zipcode



Min of price and Max of price by grade



# THANK YOU!

Ferza Reyaldi - DAI 003

Linkedin: <https://www.linkedin.com/in/freyaldi/>

Capstone Repository: [Here](#)

