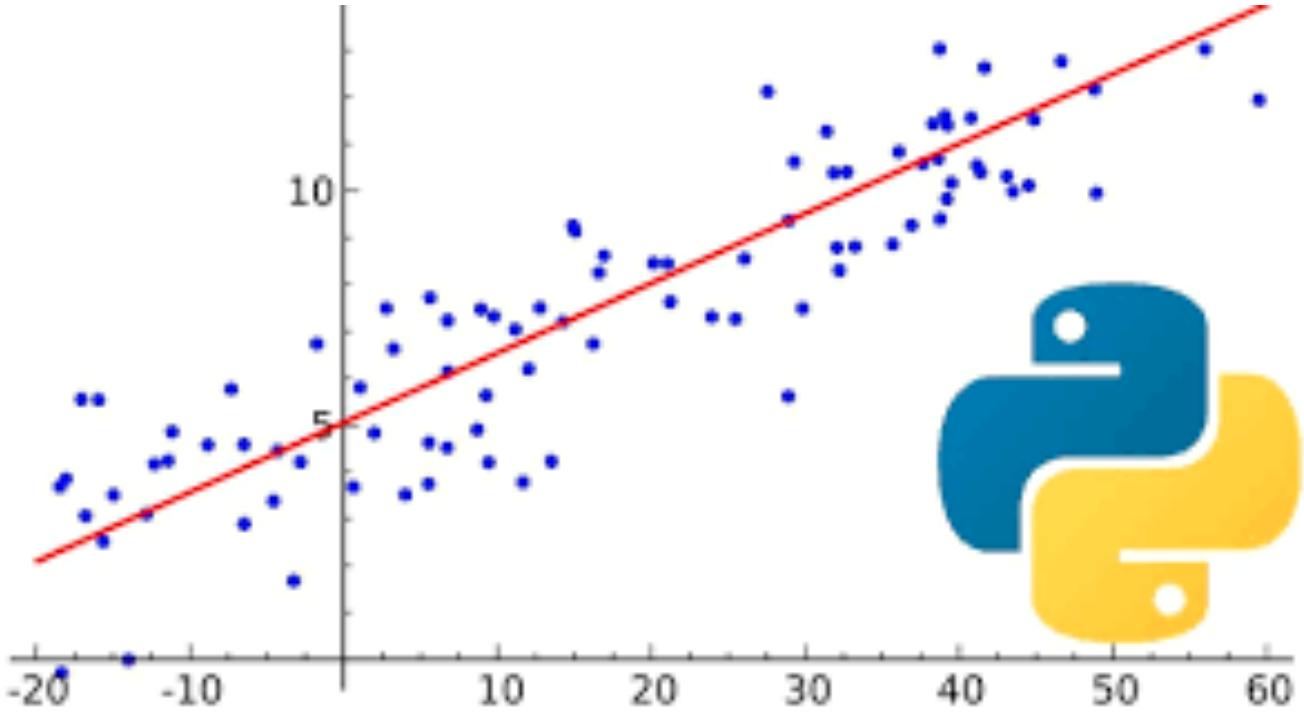
LINEAR REGRESSION USING PYTHON

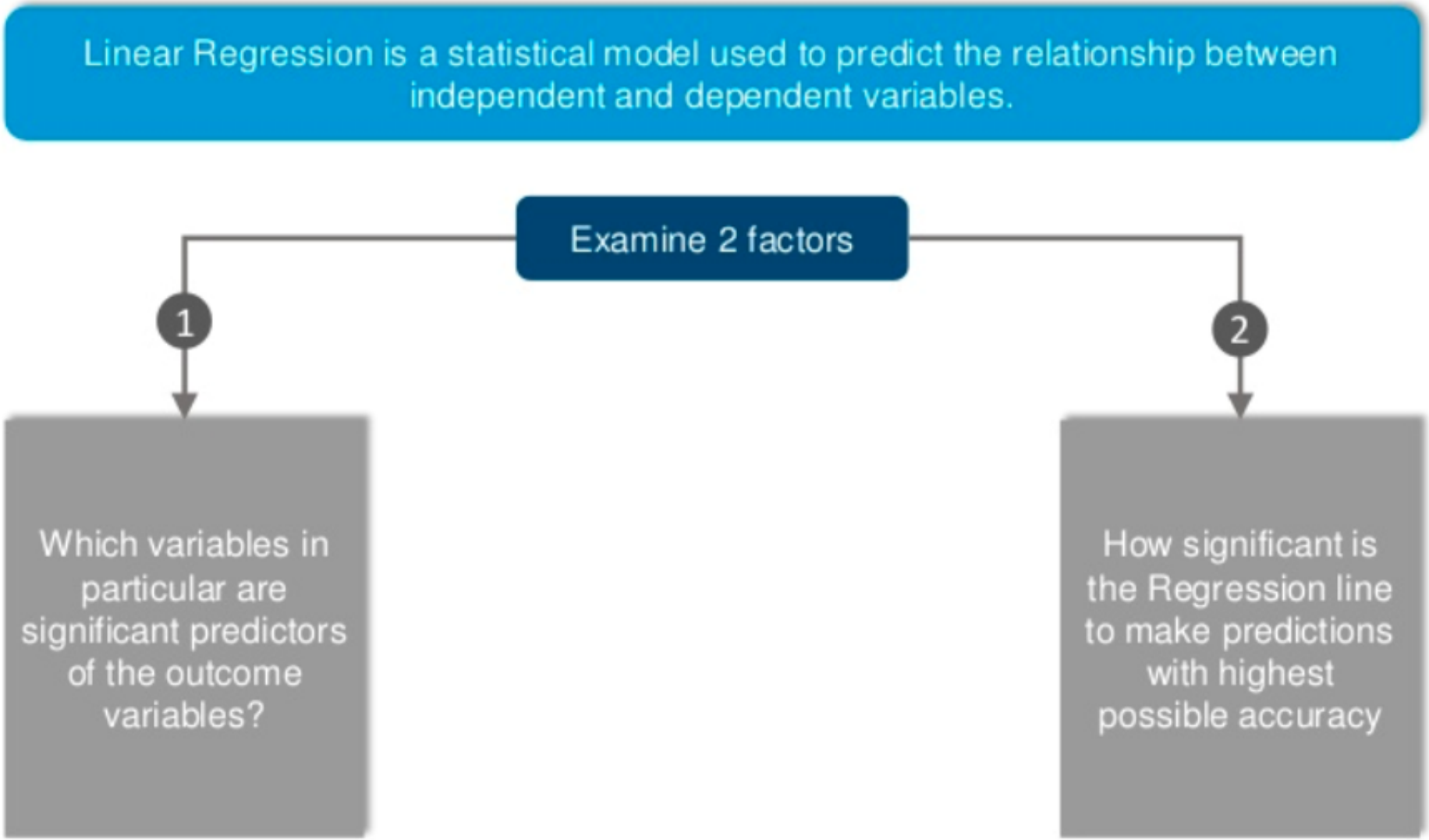


By-CHIVUKULA GAYATRI PUJA

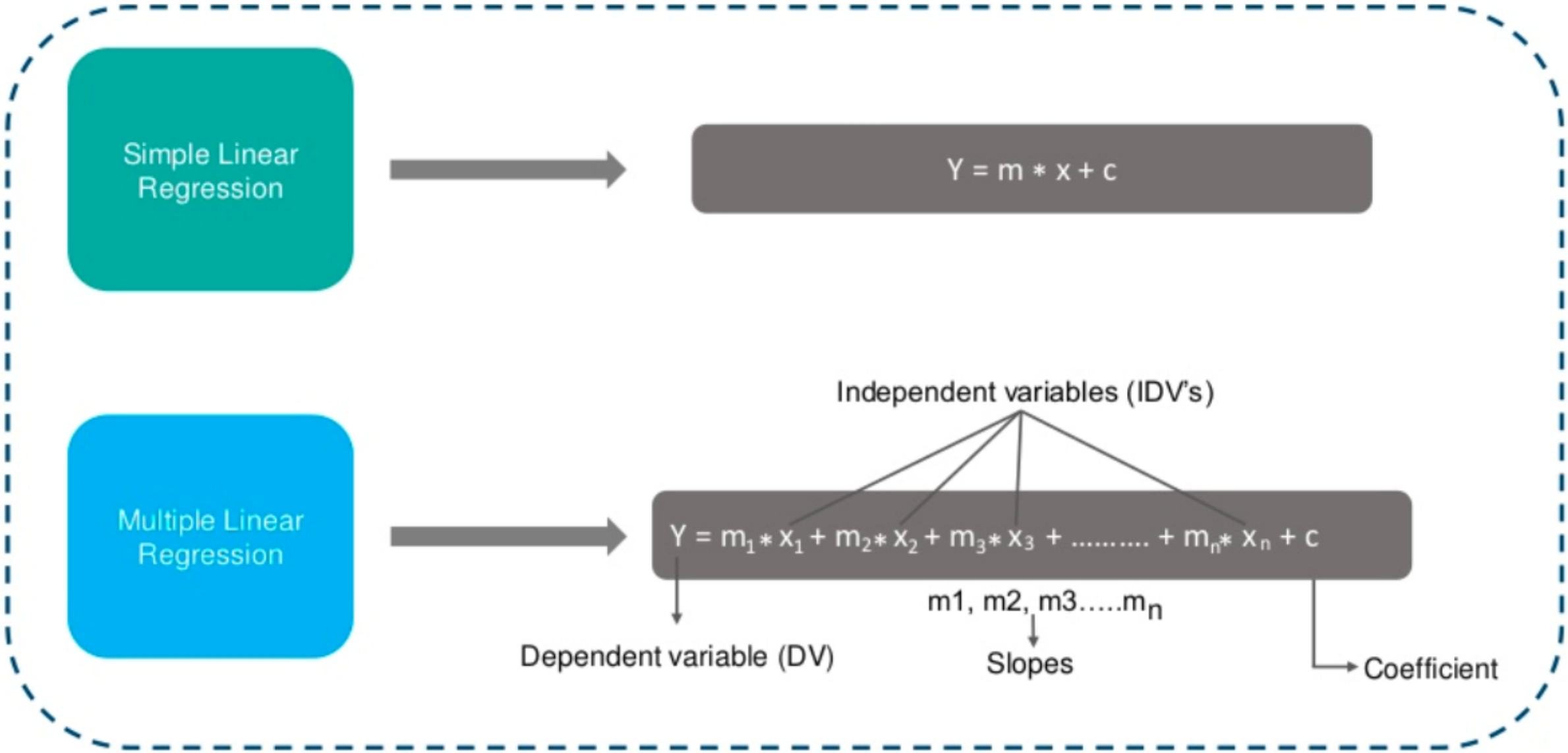
AGENDA

1. Introduction
2. Use Cases
3. Approach
4. Python Implementation
5. Evaluation Metrics
6. Predicting Prices for New Values

Introduction to Linear Regression



Mathematical model



Applications of Linear Regression

Economic Growth

Score Prediction

Product price

Use Cases

- Use **typical**, easily accessible features of cars
- Obtain **benchmark** of base car pricing
- **Cross-check** with used car listing price and car brands

Car Price Predictions for Resales

House Price Predictions

Approach : Crisp - DM

- **Business Understanding** – Predicting prices & helping companies purchase second hand cars
- **Data Understanding** – Trying to understand Independent & dependent variables & EDA
- **Data Preparation** – Data Wrangling
- **Model Building** – Fit the model by choosing the right algorithm
- **Model Evaluation** – Analyzing the model
- **Deployment**

Python Implementation (Live Coding Session)

From Data Understanding to Model Evaluation

Libraries Used :

Data Handling –
Pandas

Exploratory Data Analysis –
Pandas Profiling

Model Fitting –
Sklearn.linear_model

Evaluation Metrics –
Sklearn.Metrics

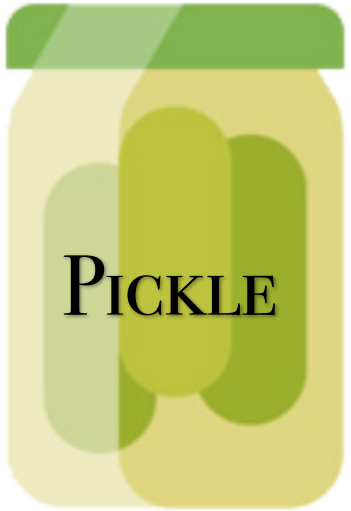
EVALUATION METRICS

Metrics library is used to calculate the Error Rate

MSE : The Mean Squared Error or Mean Squared Deviation of an estimator measures the average of error squares i.e. the average squared difference between the estimated values and true value. It is a risk function, corresponding to the expected value of the squared error loss. It is always non – negative and values close to zero are better. The MSE is the second moment of the error (about the origin) and thus incorporates both the variance of the estimator and its bias.

MAE : Mean absolute error is a measure of errors between paired observations expressing the same phenomenon. Examples of Y versus X include comparisons of predicted versus observed, subsequent time versus initial time, and one technique of measurement versus an alternative technique of measurement.

RMSE : The Root Mean Squared Error, this can be calculated by using the sqrt() math function on the mean squared error calculated using the mean_squared_error() scikit-learn function. Running the example calculates the root mean squared error.



Tech Stack

Web Application

Year	int64
Selling_Price	float64
Present_Price	float64
Kms_Driven	int64
Owner	int64
Fuel_Type_CNG	uint8
Fuel_Type_Diesel	uint8
Fuel_Type_Petrol	uint8
Seller_Type_Dealer	uint8
Seller_Type_Individual	uint8
Transmission_Automatic	uint8
Transmission_Manual	uint8

Lets predict for
[3,12,20000,2,0,1,0,1,0,1,0]

The predicted car price is

10.45

The numeric is in lakhs


thank you