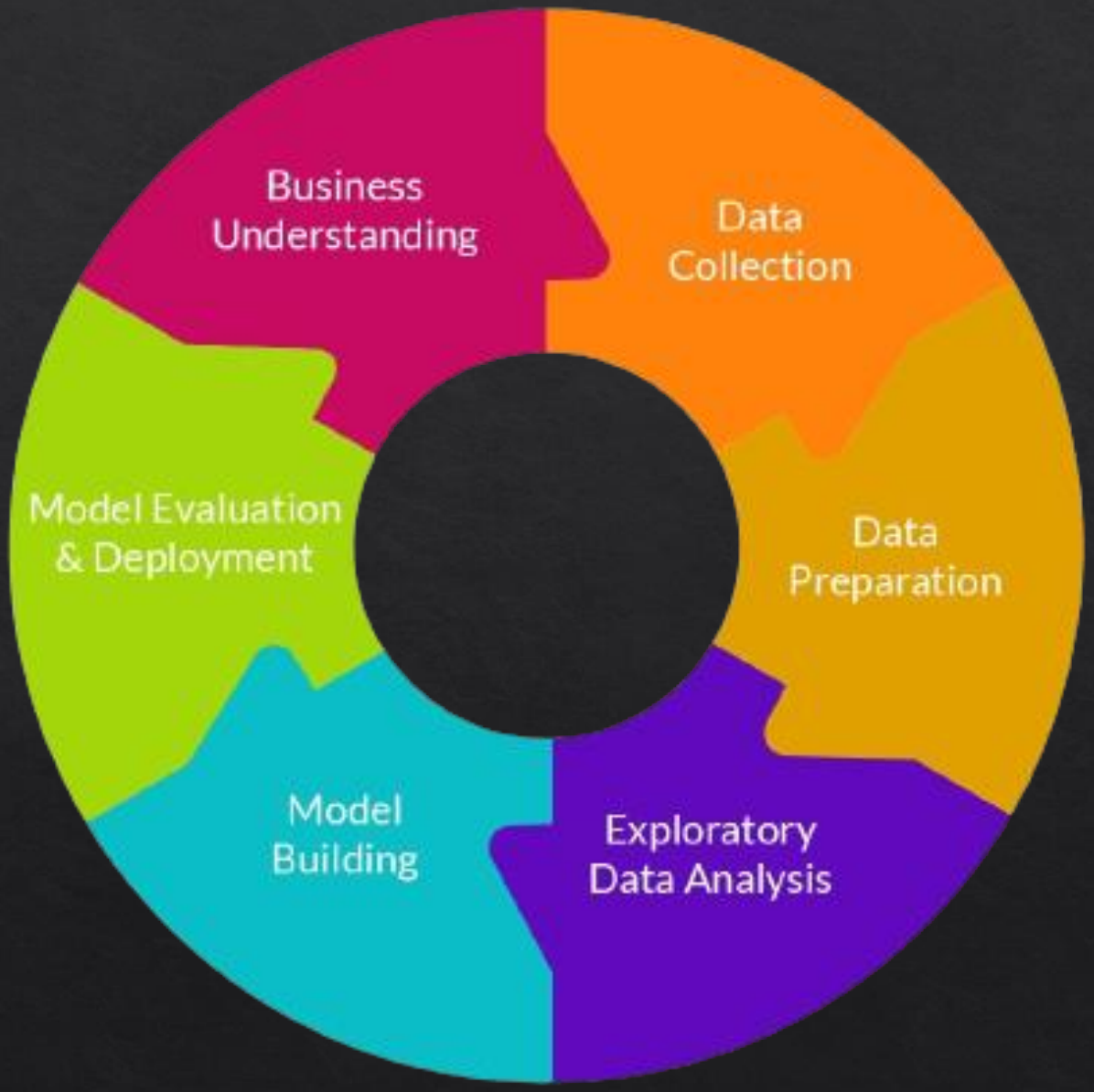
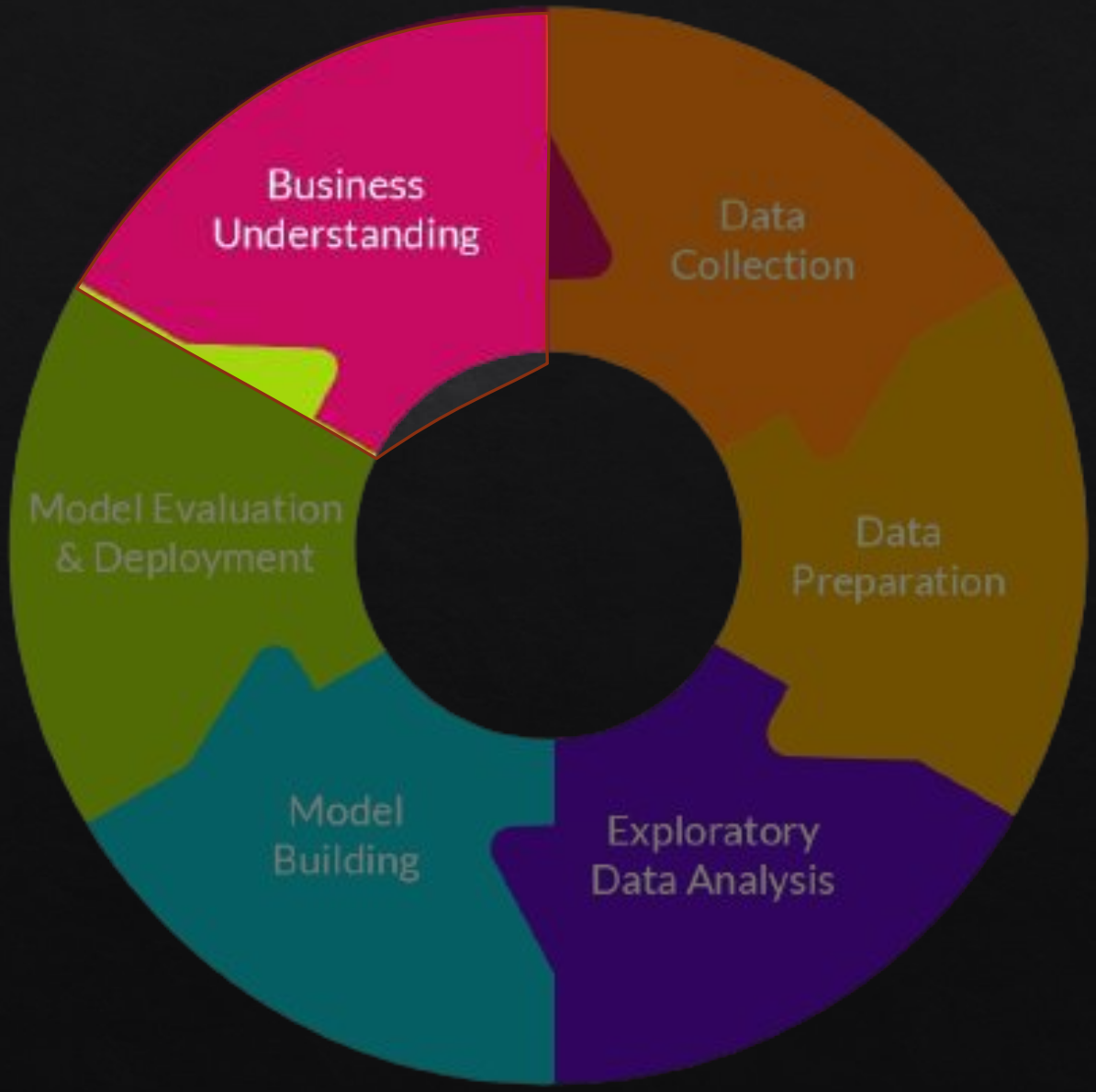
DATA SCIENCE LIFE CYCLE



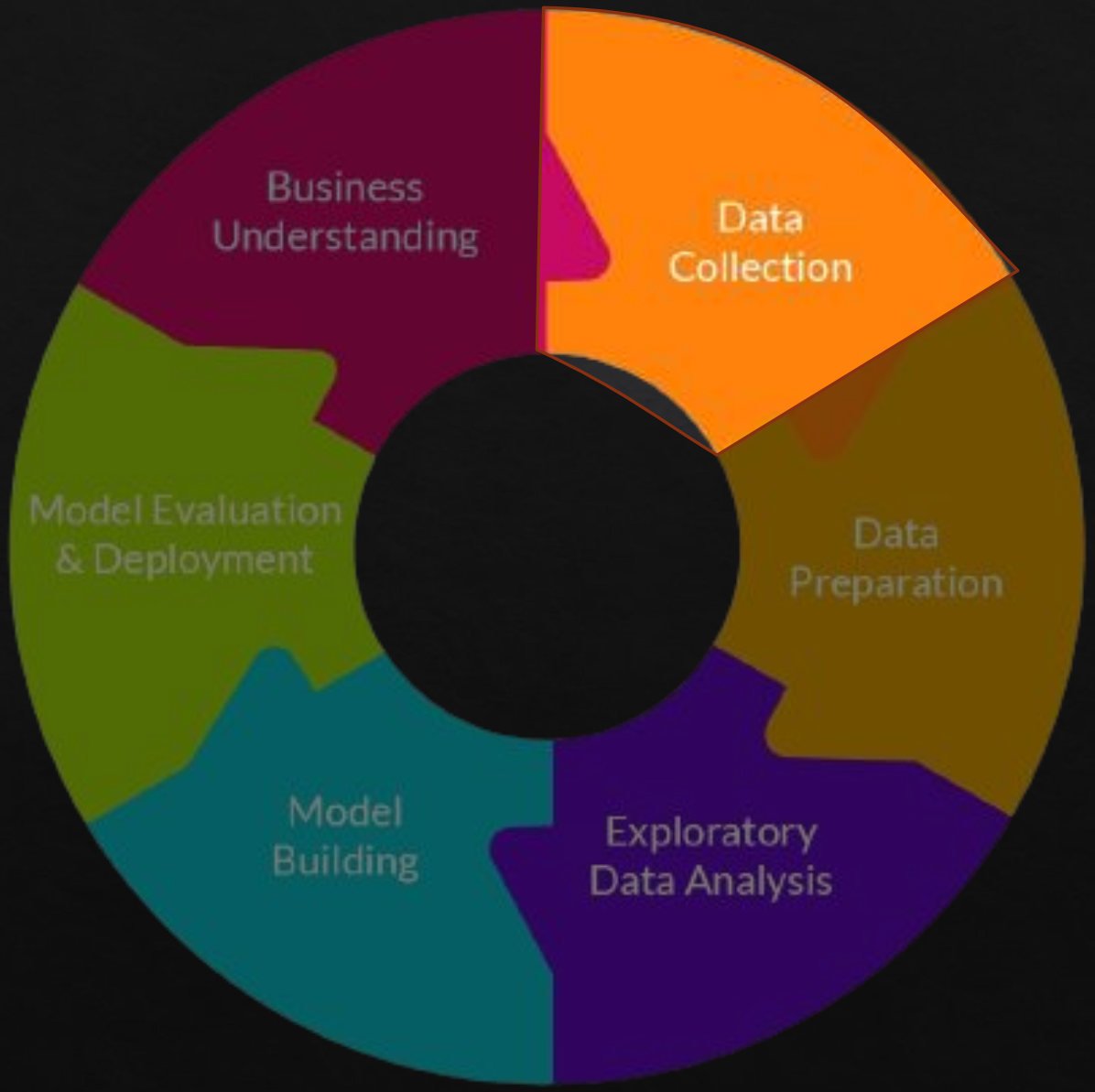
DATA SCIENCE LIFE CYCLE



PROBLEM STATEMENT AND OBJECTIVE

- The business is facing a significant inventory cost loss due to a mismatch between demand and supply across its warehouses in the country. The company's higher management is seeking to optimize the supply quantity in each warehouse to balance the demand and supply and reduce the inventory cost.
- Objective is to develop an effective strategy to optimize the supply quantity in each warehouse and minimize the inventory cost while meeting the customer demand.

DATA SCIENCE LIFE CYCLE



Warehouse_id	Competitor_in_mkt	workers_num
WH_Manager_ID	retail_shop_num	wh_est_year
Location_type	wh_owner_type	storage_issue_reported_l3m
WH_capacity_size	distributor_num	temp_reg_mach
Zone	flood_impacted	approved_wh_govt_certificate
WH_regional_zone	Flood_proof	wh_breakdown_l3m
num_refill_req_l3m	electric_supply	govt_check_l3m
transport_issue_l1y	dist_from_hub	**product_wg_ton** (Target column)

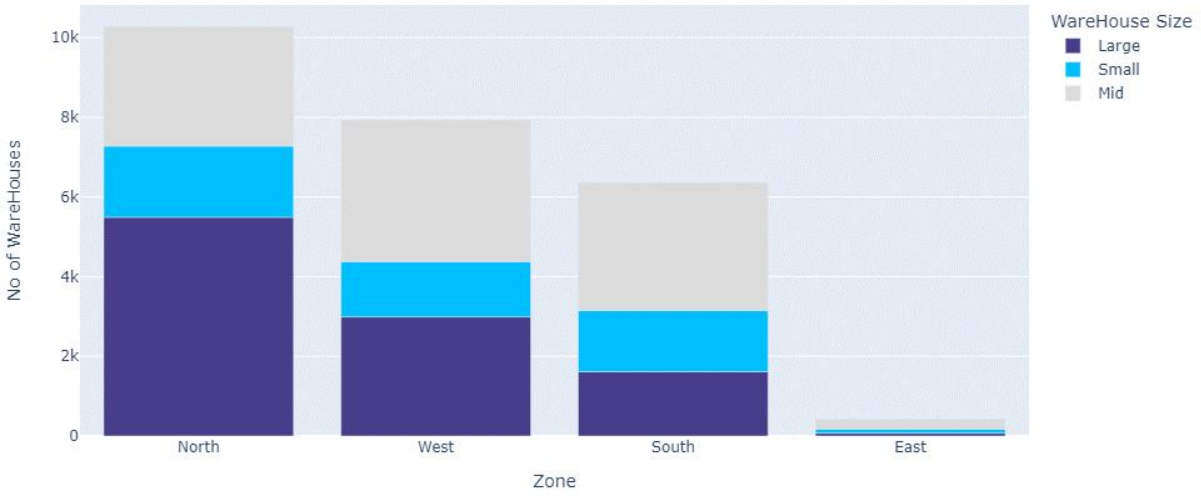
DATA SETS	NO OF RECORDS	NO OF COLUMNS	NO OF DUPLICATE RECORDS	TOTAL NULL VALUES	TOTAL NOISES IN DATA	UNIQUE WARE HOUSE ID VALUES
FMCG DATA	25000	24	0	13779	0	25000

	Missing_Number	Missing_Percent
wh_est_year	11881	0.47524
workers_num	990	0.03960
approved_wh_govt_certificate	908	0.03632

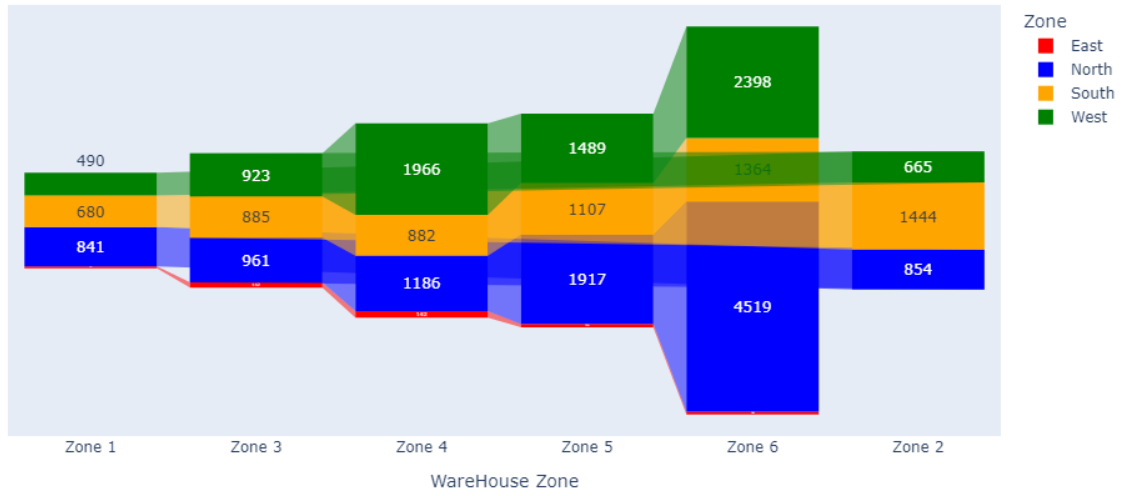
DATA SCIENCE LIFE CYCLE

NUMBER OF WARE HOUSES

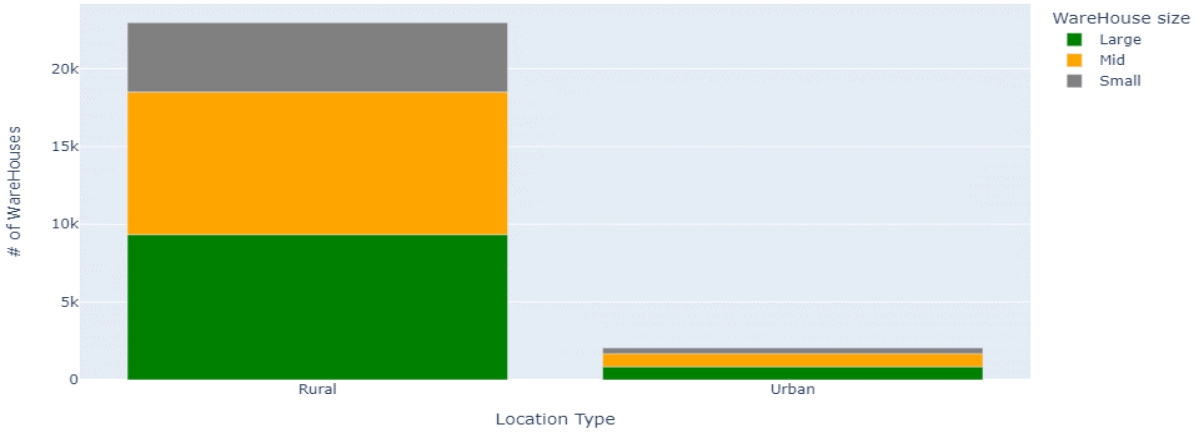
No of WareHouses in each Zone according to size



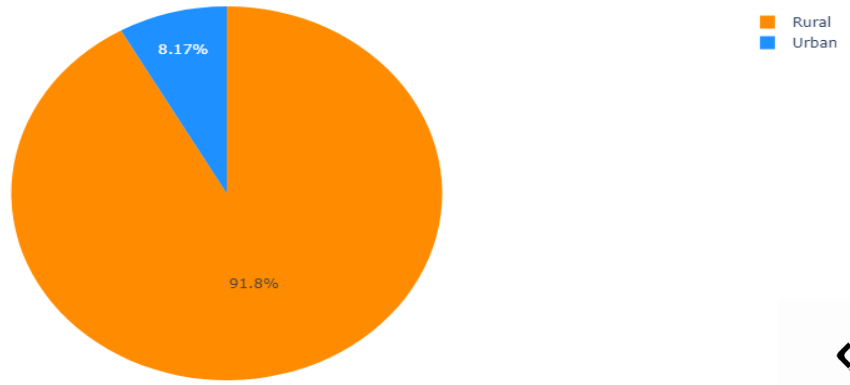
No of WareHouses refilled in different zones in last 3 months



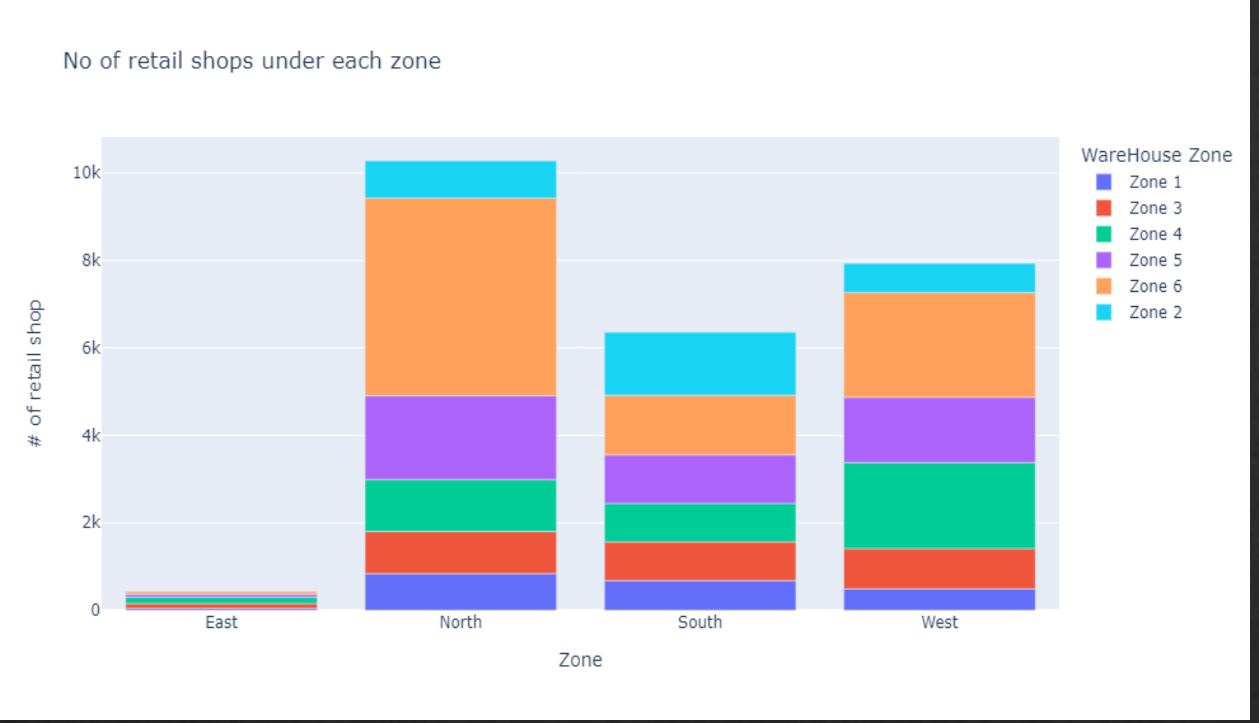
No of WareHouses in different location

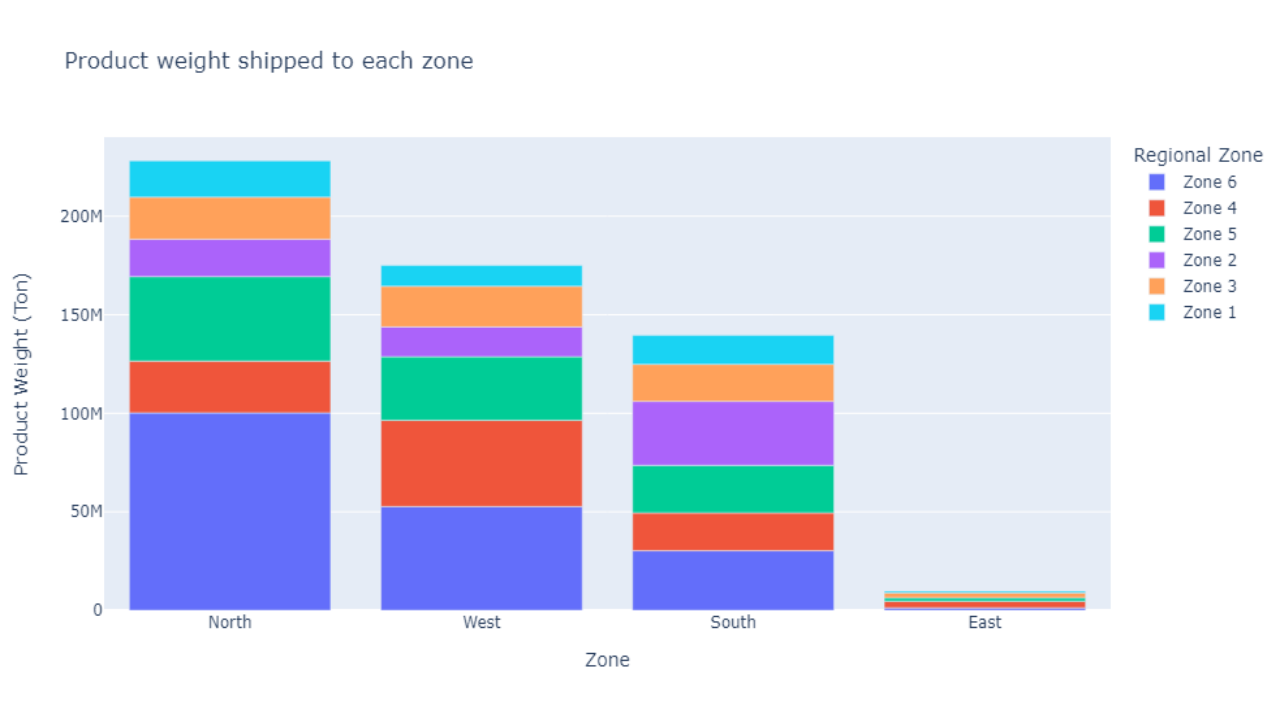


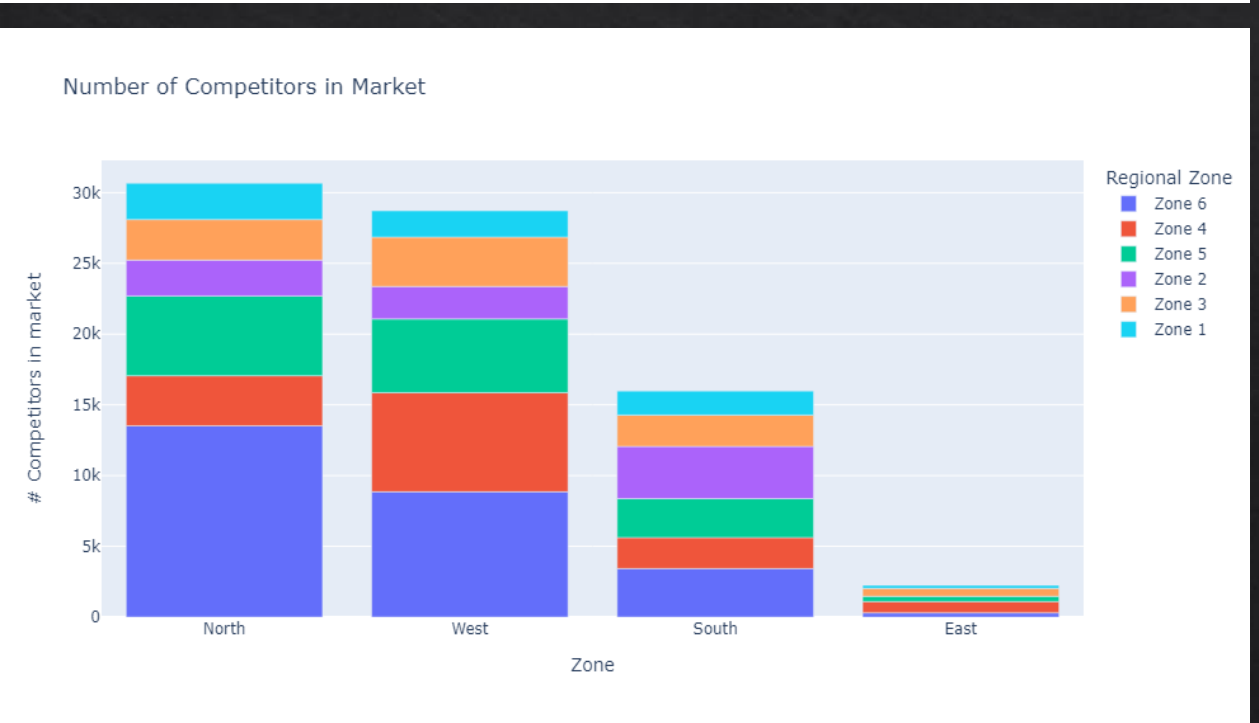
Number of retail shops

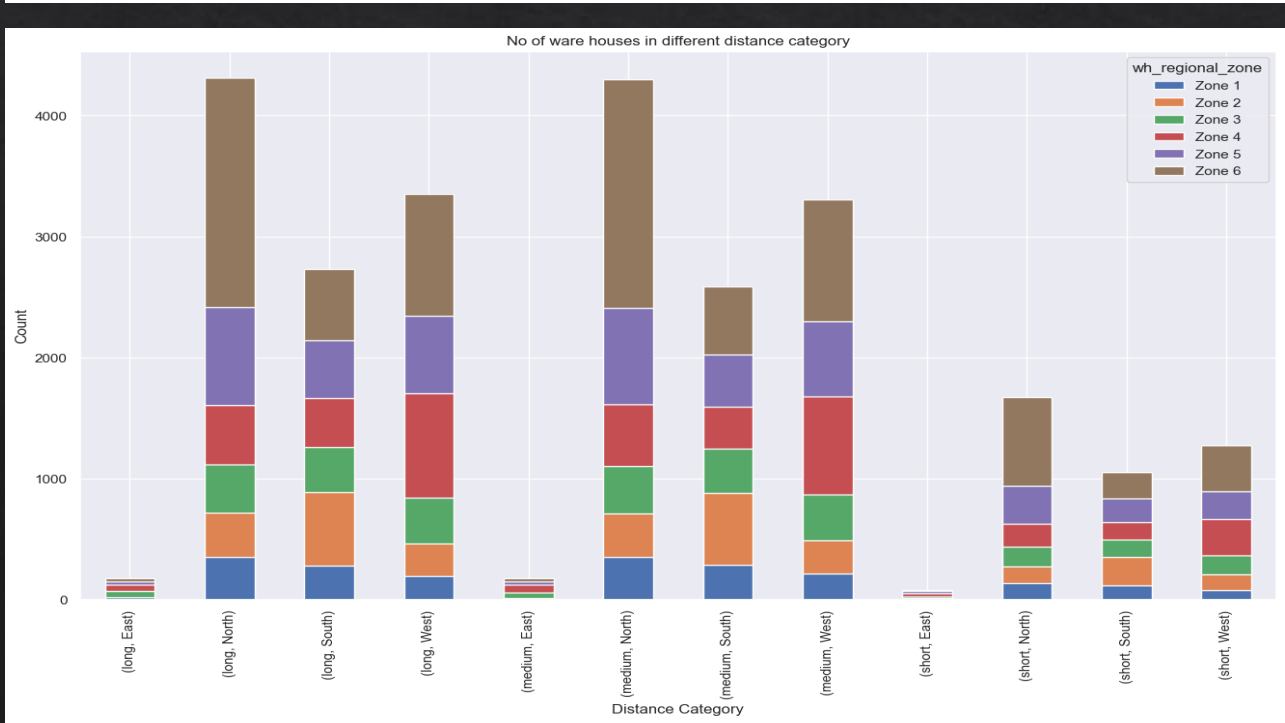


Learnbay

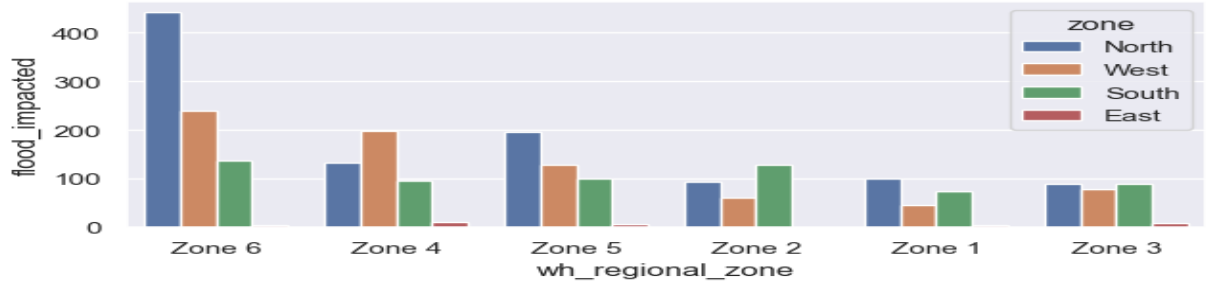

No of retail shops under each zone
# of retail shop
10k
8k
6k
4k
2k
0
East
North
South
West
Zone
WareHouse Zone
Zone 1
Zone 3
Zone 4
Zone 5
Zone 6
Zone 2

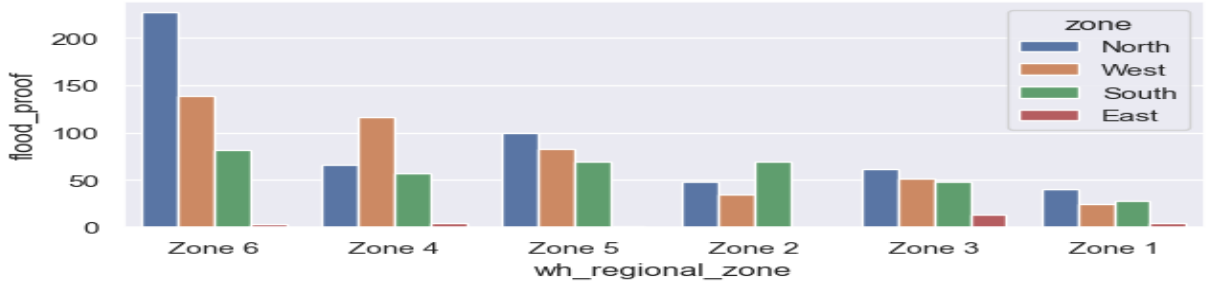

Product weight shipped to each zone
Product Weight (Ton)
200M
150M
100M
50M
0
North
West
South
East
Zone
Regional Zone
Zone 6
Zone 4
Zone 5
Zone 2
Zone 3
Zone 1

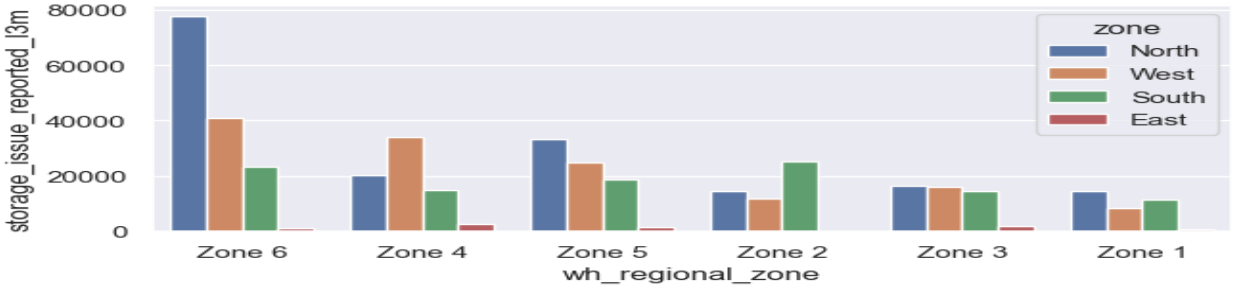

Number of Competitors in Market
Competitors in market
30k
25k
20k
15k
10k
5k
0
North
West
South
East
Zone
Regional Zone
Zone 6
Zone 4
Zone 5
Zone 2
Zone 3
Zone 1

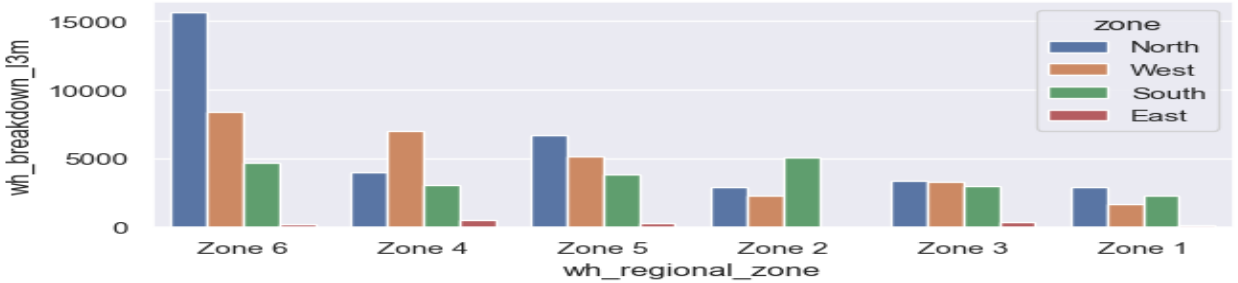

No of ware houses in different distance category
Count
4000
3000
2000
1000
0
(long, East)
(long, North)
(long, South)
(long, West)
(medium, East)
(medium, North)
(medium, South)
(medium, West)
(short, East)
(short, North)
(short, South)
(short, West)
Distance Category
wh_regional_zone
Zone 1
Zone 2
Zone 3
Zone 4
Zone 5
Zone 6

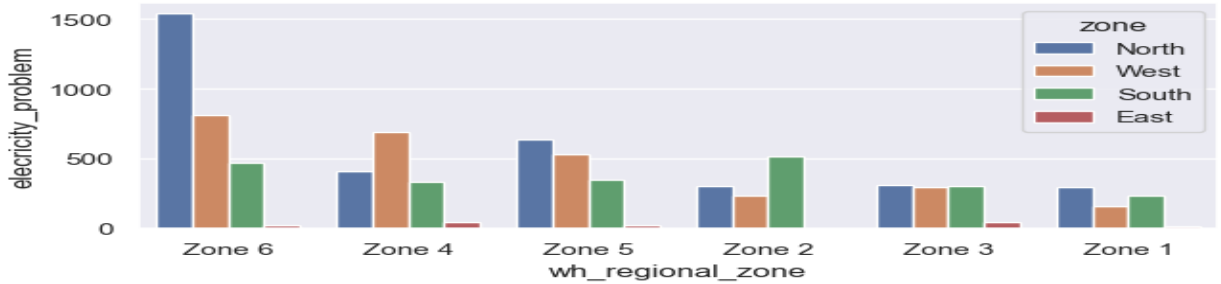
ISSUES IN WARE HOUSE

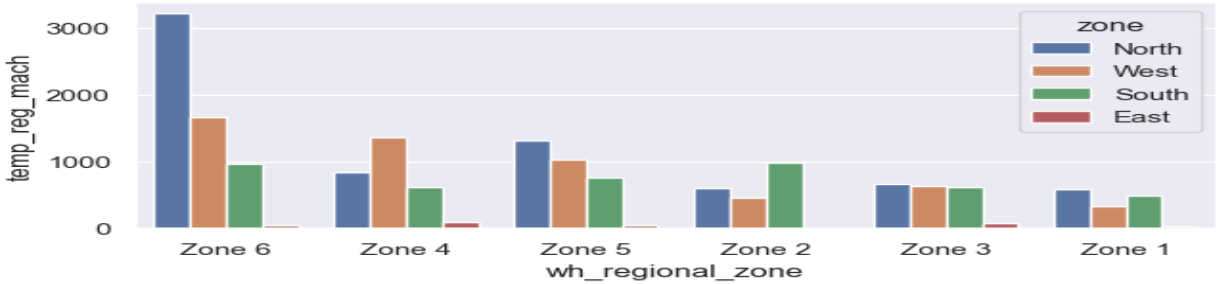


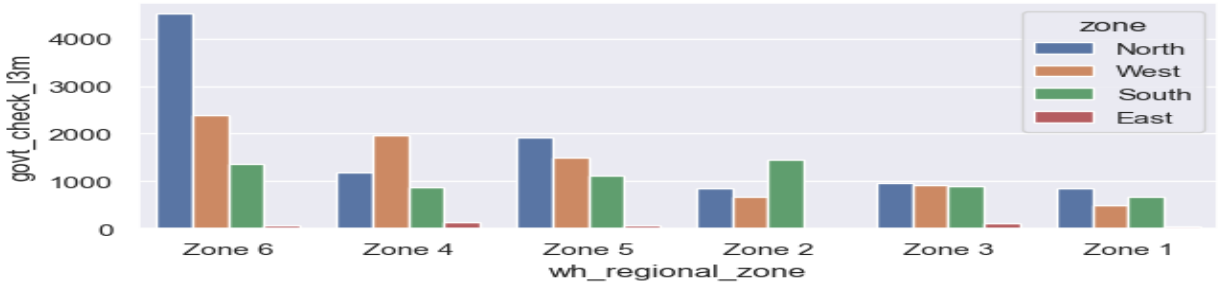


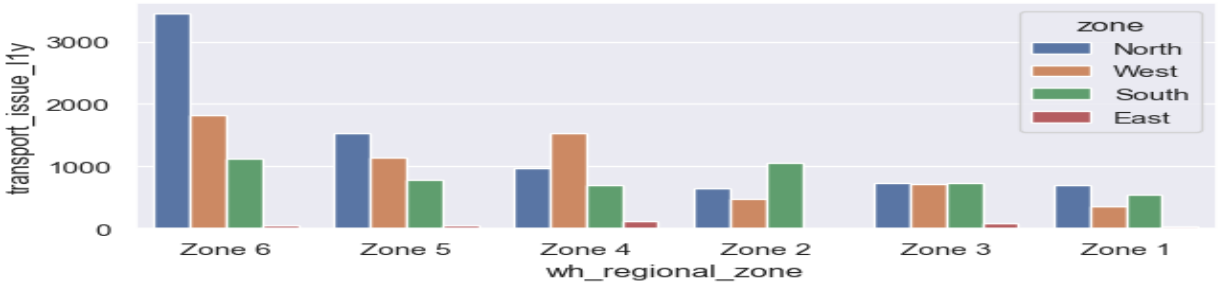






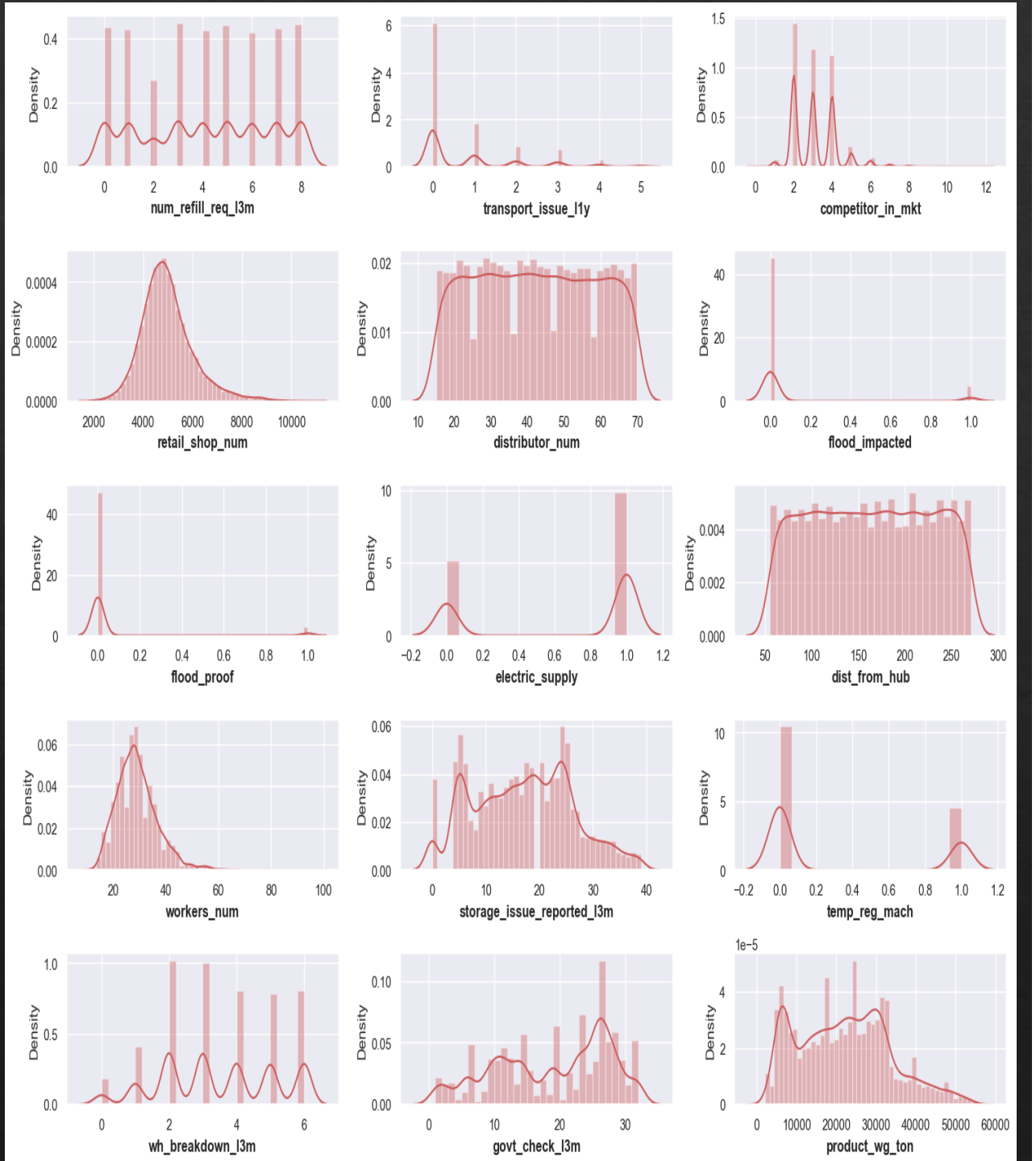


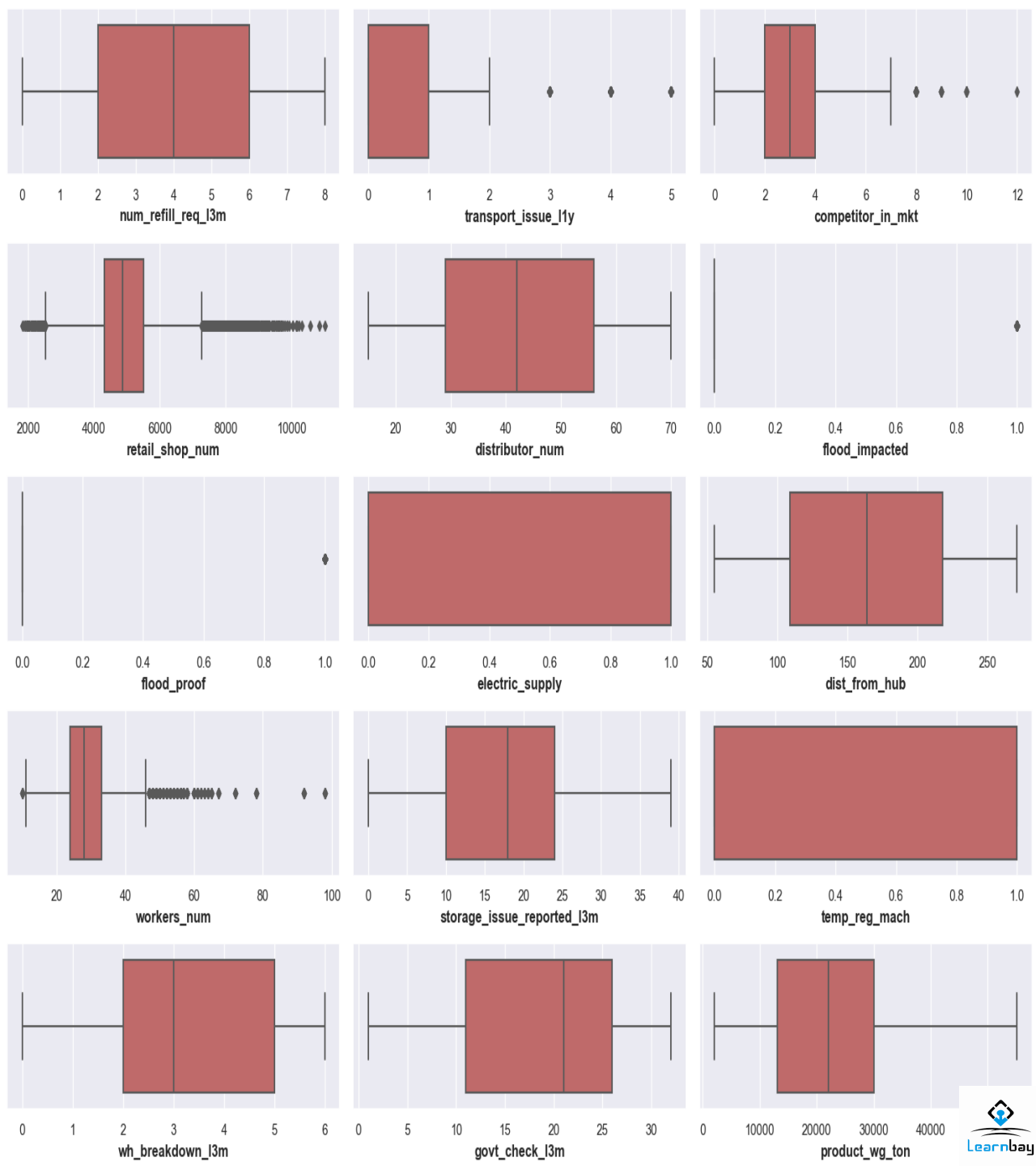




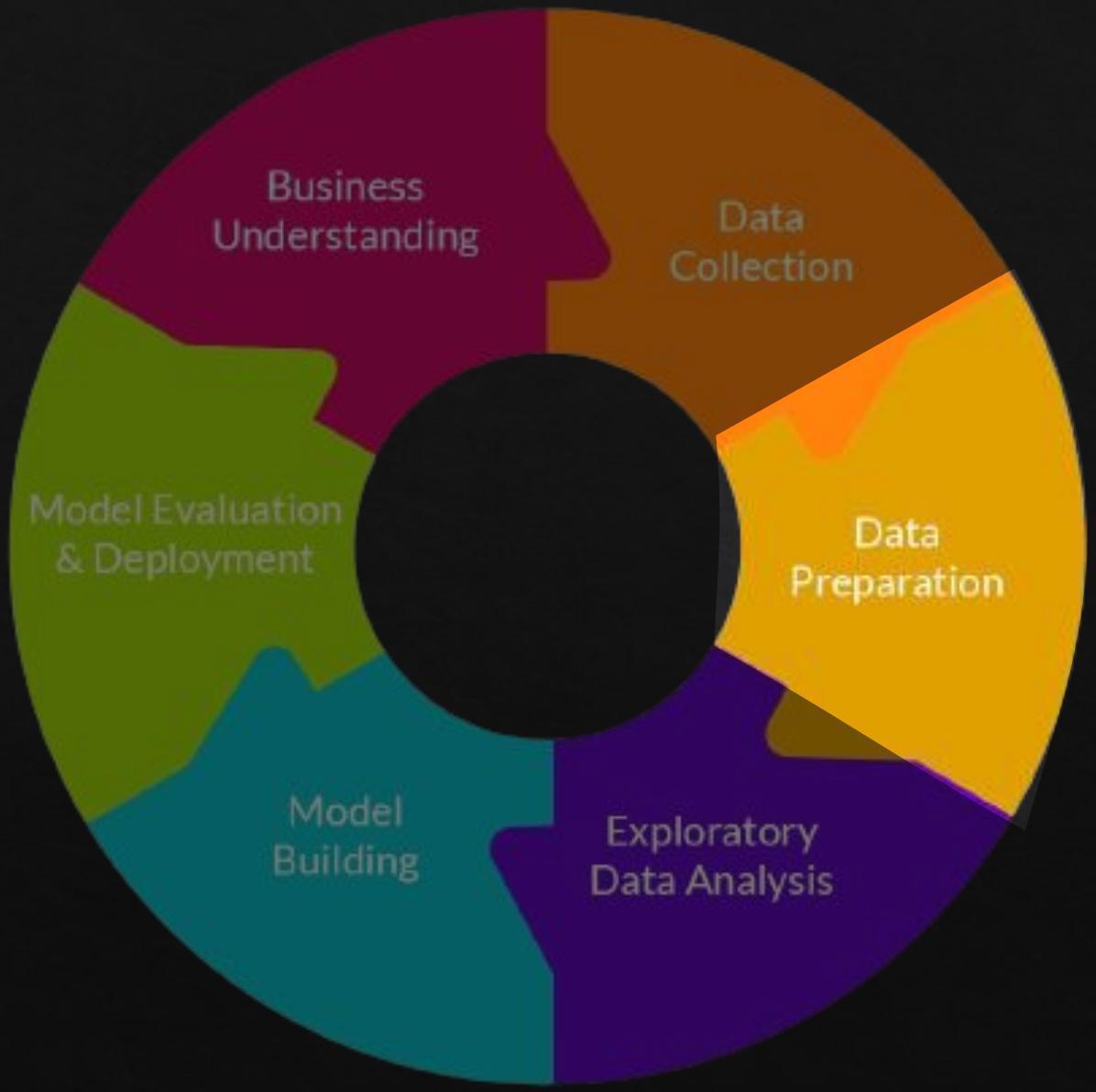
SL NO	TITLE	1ST	2ND	3RD	4TH	5TH
1	Number of warehouses in each zone	North – Zone 6	West – Zone 6	West – Zone 4	North – Zone 5	West – Zone 5
2	Number of warehouses refilled in last 3 months	North – Zone 6	West – Zone 6	West – Zone 4	North – Zone 5	West – Zone 5
3	Number of retail shops in each zone	North – Zone 6	West – Zone 6	West – Zone 4	North – Zone 5	West – Zone 5
4	Product weight shipped to each zone	North – Zone 6	West – Zone 6	West – Zone 4	North – Zone 5	South – Zone 2
5	Number of Competitors in Market	North – Zone 6	West – Zone 6	West – Zone 4	North – Zone 5	West – Zone 5
6	Number of distributors in each zone	North – Zone 6	West – Zone 6	West – Zone 4	North – Zone 5	West – Zone 5
7	Issues in warehouse	North – Zone 6	West – Zone 6	West – Zone 4	North – Zone 5	West – Zone 5

SL NO	TITLE	1ST	2ND	3RD	4TH	5TH
1	Number of warehouses in each zone	East – Zone 4	East – Zone 3	East – Zone 5	East – Zone 6	East – Zone 1
2	Number of warehouses refilled in last 3 months	East – Zone 4	East – Zone 3	East – Zone 5	East – Zone 6	East – Zone 1
3	Number of retail shops in each zone	East – Zone 4	East – Zone 3	East – Zone 5	East – Zone 6	East – Zone 1
4	Product weight shipped to each zone	East – Zone 4	East – Zone 3	East – Zone 5	East – Zone 6	East – Zone 1
5	Number of Competitors in Market	East – Zone 4	East – Zone 3	East – Zone 5	East – Zone 6	East – Zone 1
6	Number of distributors in each zone	East – Zone 4	East – Zone 3	East – Zone 5	East – Zone 6	East – Zone 1
7	Issues in warehouse	East – Zone 4	East – Zone 3	East – Zone 5	East – Zone 6	East – Zone 1


Density
num_refill_req_l3m
transport_issue_l1y
competitor_in_mkt
retail_shop_num
distributor_num
flood_impacted
flood_proof
electric_supply
dist_from_hub
workers_num
storage_issue_reported_l3m
temp_reg_mach
wh_breakdown_l3m
govt_check_l3m
product_wg_ton
1e−5


num_refill_req_l3m
transport_issue_l1y
competitor_in_mkt
retail_shop_num
distributor_num
flood_impacted
flood_proof
electric_supply
dist_from_hub
workers_num
storage_issue_reported_l3m
temp_reg_mach
wh_breakdown_l3m
govt_check_l3m
product_wg_ton

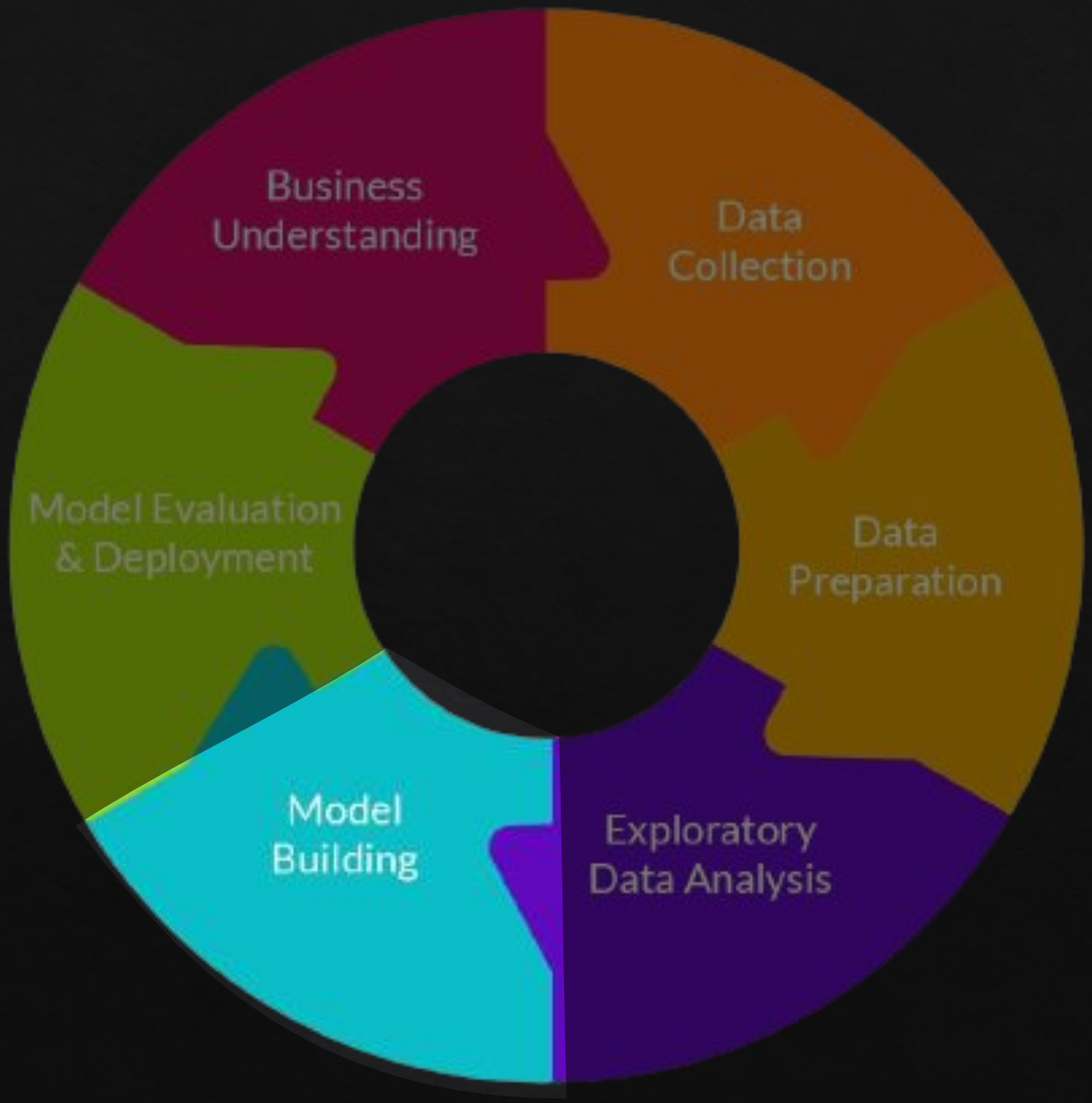
DATA SCIENCE LIFE CYCLE



DATA IMPUTATION AND TRANSFORMATION

- Mapping
- One hot encoding
- Robust Scaler
- KNN imputation
- VIF
- Lasso Regression
- Pipeline

DATA SCIENCE LIFE CYCLE

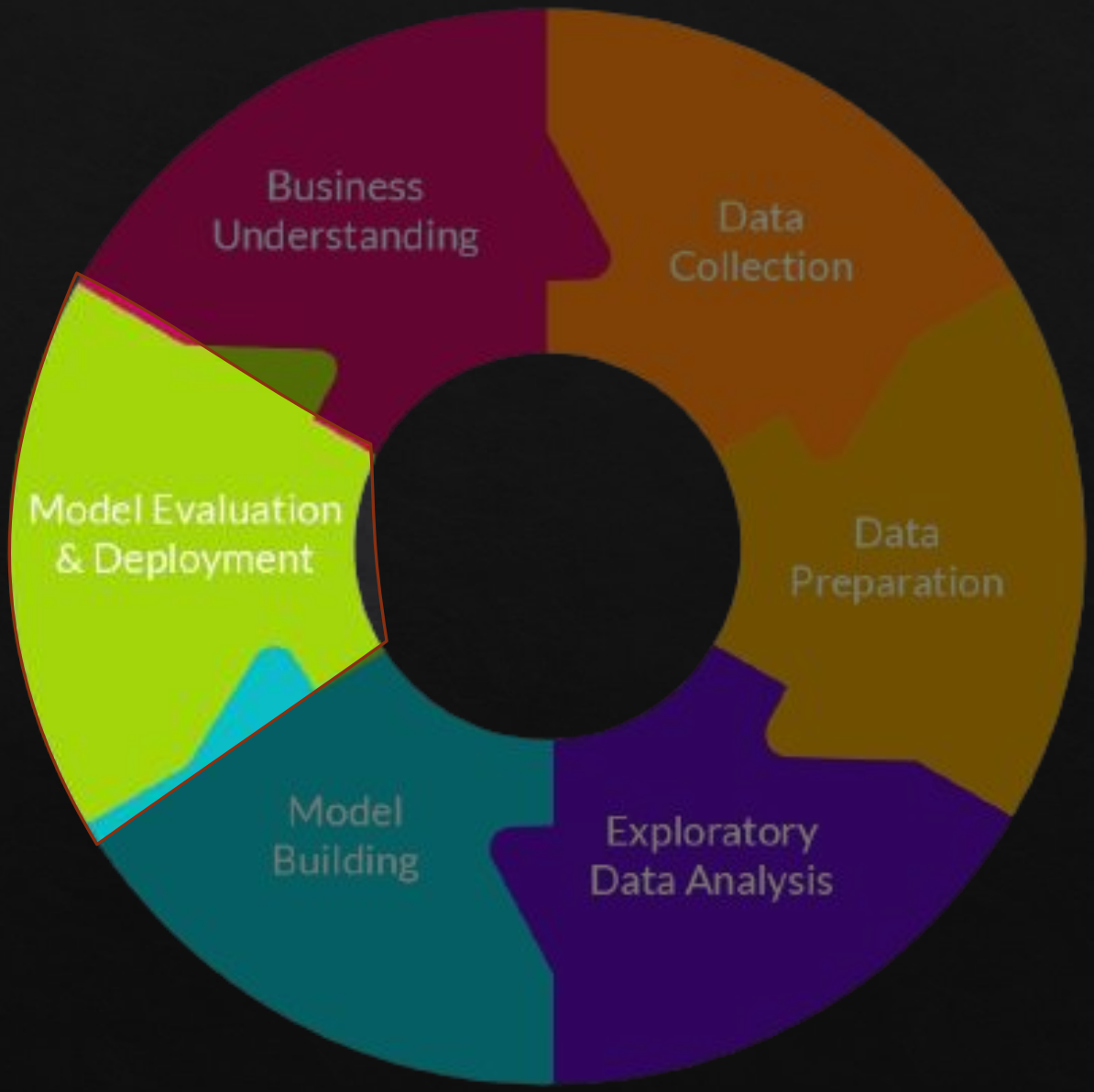


SELECTING ALGORITHMS

- Initialize Default Models in a dictionary
- Model result
- Predictions

- Linear Regression
- Random Forest
- Gradient Boosting

DATA SCIENCE LIFE CYCLE



MODEL EVALUATION AND DEPLOYMENT

- Cross Validation
- Saving model
- Deployment

SL NO	MODEL	PERFORMANCE			
		TRAIN ERROR	TEST ERROR	TRAIN SCORE	TEST SCORE
1	Linear Regression	1716.63	1771.98	97.8%	97.7%
2	Random Forest	360.24	965.81	99.9%	99.3%
3	Gradient Boosting	915.85	947.68	99.4%	99.3%

CONCLUSION

The competitive market

- Differentiate your product
- Focus on quality
- Offer competitive pricing
- Invest in marketing
- Build strong distribution channels

Issues in the warehouse

- Problems like floods, storage issue, WH breakdown, electricity issue, temperature regulator issue, will have a significant impact on businesses
- Number of time government Officers visited the warehouse in last 3 months may be because of complaints received from the dissatisfied customers

Distance Factor

- Shipping costs
- Delivery time
- Supply chain complexity

Optimum weight of the product

Gradient Boosting algorithm works as the best model to predict the optimum weight of the product to be shipped to the warehouse

FUTURE SCOPE

TRAINING PIPELINE