

# Recommender System





# Hello!

Zia Khan

[zia@thedevmasters.com](mailto:zia@thedevmasters.com)



# CRISP-DM

1

Business Objective

2

Data Understanding

3

Data Preparation

4

Modeling

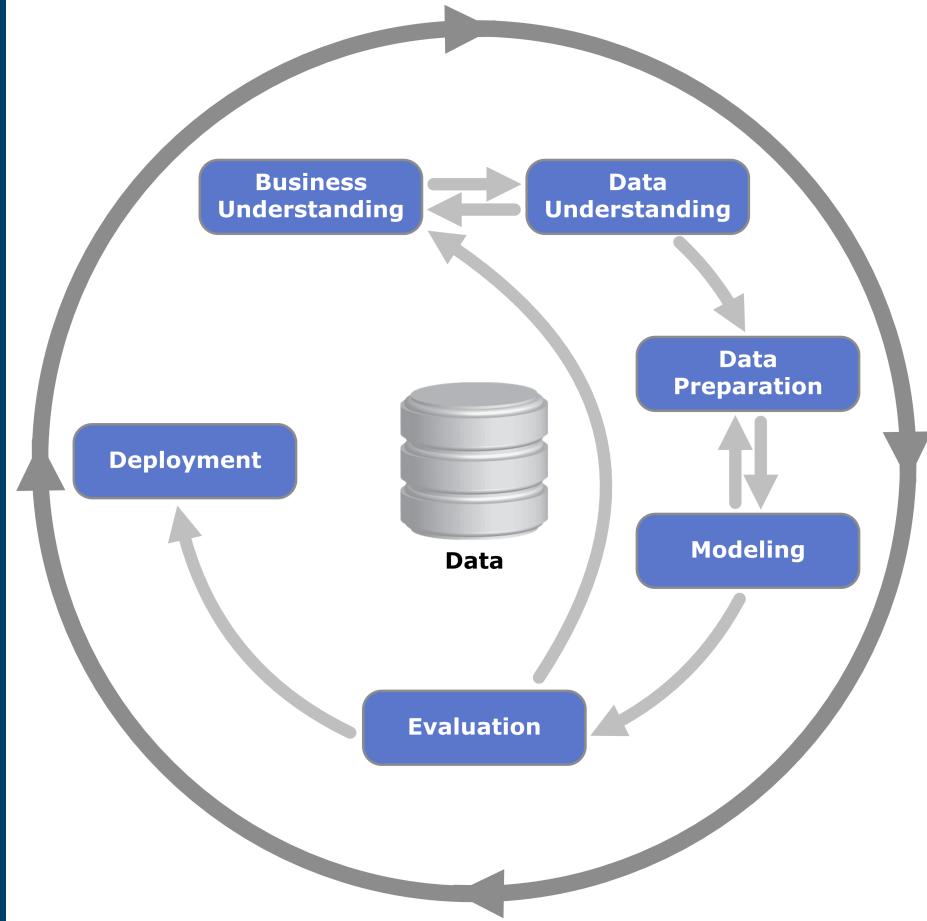
5



# CRISP-DM

## Overview

- Business Understanding
- Data Understanding
- Data Preparation
- Modeling
- Evaluation
- Deployment



**CRISP-DM**

**1**

**Business Objective**

**2**

**Data Understanding**

**3**

**Data Preparation**

**4**

**Modeling**

**5**



# 2. Build a recommender system

The goal of the project is to help the company generate incremental sales by implementing a recommender system.



**CRISP-DM**

**1**

**Business Objective**

**2**

**Data Understanding**

**3**

**Data Preparation**

**4**

**Modeling**

**5**



# Data Understanding

	Customers.id	Customers.fname	Customers.lname	Customers.company	Customers.create_date	Customers.status	Customers.mailing	Customers.reminders
0	797	Christy	Dill	Company0	1426018724	NaN	NaN	NaN
1	3	John	Smith	Company1	1386089139	NaN	NaN	NaN
2	3	John	Smith	Company1	1386089139	NaN	NaN	NaN
3	4	James	Anderson	NaN	1386780263	NaN	NaN	NaN
4	5	Abraham	Pollak	Company3	1386861599	0.0	0.0	NaN
5	5	Abraham	Pollak	Company3	1386861599	0.0	0.0	NaN
6	5	Abraham	Pollak	Company3	1386861599	0.0	0.0	NaN
7	7	peggy	thompson	NaN	1388155947	NaN	NaN	NaN
8	8	Randy	Pruss	NaN	1389303216	NaN	NaN	NaN
9	10	Tommy	Smith	NaN	1390509516	NaN	NaN	NaN
10	11	Mark	Tremble	NaN	1390573155	NaN	NaN	NaN
11	12	Emely	Cooke	NaN	1390613314	NaN	NaN	NaN
12	13	george	mcmillin	NaN	1390856039	NaN	NaN	NaN
13	13	george	mcmillin	NaN	1390856039	NaN	NaN	NaN
14	13	george	mcmillin	NaN	1390856039	NaN	NaN	NaN

15 rows x 181 columns



# Data Understanding

Products.hcpcs	906
Products.case_qty	1556
Products.markup	4194
Products.override_markup	1274
Products.notes	3748
Products.import_flags	1408
Products.map_price	4072
Products.features_title	4192
Products.warranty	3531
Products.hygienic	2114
Products.default_quantity	4194
Products.user_size	3578
Products.assembly	3578
Products.installation	3578
Products.shipping_length	196
Products.shipping_width	196
Products.shipping_height	196
Products.shipping_weight	3535
Products.handling_time	3196
Products.rotation_link	3994
Products.google_shopping_label	3807
Products.product_option	3881
Products.size	2751
Products.material	3392
Products.arm_style	4039
Products.leg_style	4194
Products.seat_size	4032
Products.family_id	325
Products.saved_status	264
Products.freight_cost	4194

Customers.id	0
Customers.fname	0
Customers.lname	0
Customers.company	3467
Customers.create_date	0
Customers.status	4093
Customers.mailing	2080
Customers.reminders	4194
Customers.tax_exempt	4189
Customers.account_id	4191
Customers.sales_rep	4194
Customers.rewards	4194
Customers.profile_id	4194
Customers.last_modified	0
Customers.customer_type	423
Orders.id	0
Orders.customer_id	0
Orders.fname	0
Orders.lname	0
Orders.company	3466
Orders.order_number	0
Orders.reorder_id	4112
Orders.external_source	3829
Orders.external_id	4189
Orders.currency	0
Orders.sales_rep	4194
Orders.subtotal	0
Orders.tax	457
Orders.shipping	195
Orders.coupon_id	3312



**CRISP-DM**

**1**

**Business Objective**

**2**

**Data Understanding**

**3**

**Data Preparation**

**4**

**Modeling**

**5**



# Data Prep and Modelling

- Missing data
- Calculate find most popular product using maximum sales by volume and dollar amount
- Company with maximum purchase
- Build a popularity based recommender system
- Build a recommender system using matrix factorization
- Cosine similarity for product description





# Thanks!

Any questions?

[zia@thedevmasters.com](mailto:zia@thedevmasters.com)

