MSDS 7330
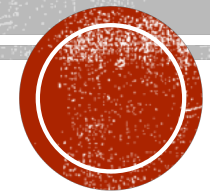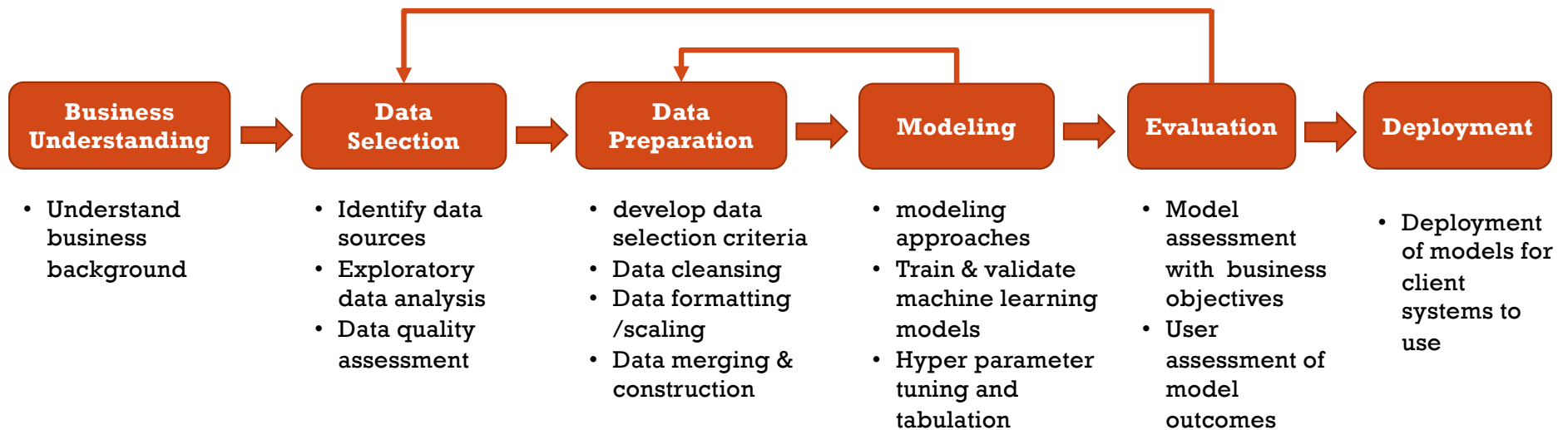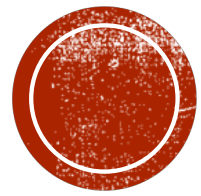# FILE ORGANIZATION AND DATABASE MANAGEMENT

Agenda

- Data Mining – Development Process

- Rest Services – Hands On Lab

- Rest Service – In Class Assignment

- API Documentation using SWAGGER

- PMML

- Homework

Typical Data Sciences : Development Process

- Iterative development process
- Model training & refinement typically carried out X sprints

Business Understanding	Data Selection	Data Preparation	Modeling	Evaluation	Deployment

- Understand business background

- Identify data sources
- Exploratory data analysis
- Data quality assessment

- develop data selection criteria
- Data cleansing
- Data formatting /scaling
- Data merging & construction

- modeling approaches
- Train & validate machine learning models
- Hyper parameter tuning and tabulation

- Model assessment with business objectives
- User assessment of model outcomes

- Deployment of models for client systems to use

HANDS-ON LAB

API DOCUMENTATION

PMML

Homework

- Using our EcommerceDB , build a recommendation rest service that takes in a customer as input and returns the item for which a promotion can be offered for that customer.

- Your recommendation model can be as simple as finding the most frequently bought item by that customer or anything else you choose.

- Turn in your HW with the codebase and a small write describing your model.