

Final Exam

DATA MODELLING AND PROBABILITY ANALYSIS

Professor HG Locklear
hlocklear@pace.edu

General

2

- ▶ This is your Final Exam and is worth 30 points.
- ▶ The Exam solution is due 8 AM Dec 23.
- ▶ The Exam Solution should be contained in a single `.ipynb` file `[Your Last Name].ipynb`
- ▶ Organize your Jupyter Notebook to clearly show the steps in your solution.
- ▶ The Exam evaluates your ability to conduct data modelling by building a data processing pipeline that involves
 - ▶ Data Cleaning
 - ▶ Data Shaping
 - ▶ Data Analysis
 - ▶ Model Building
 - ▶ Model Analysis

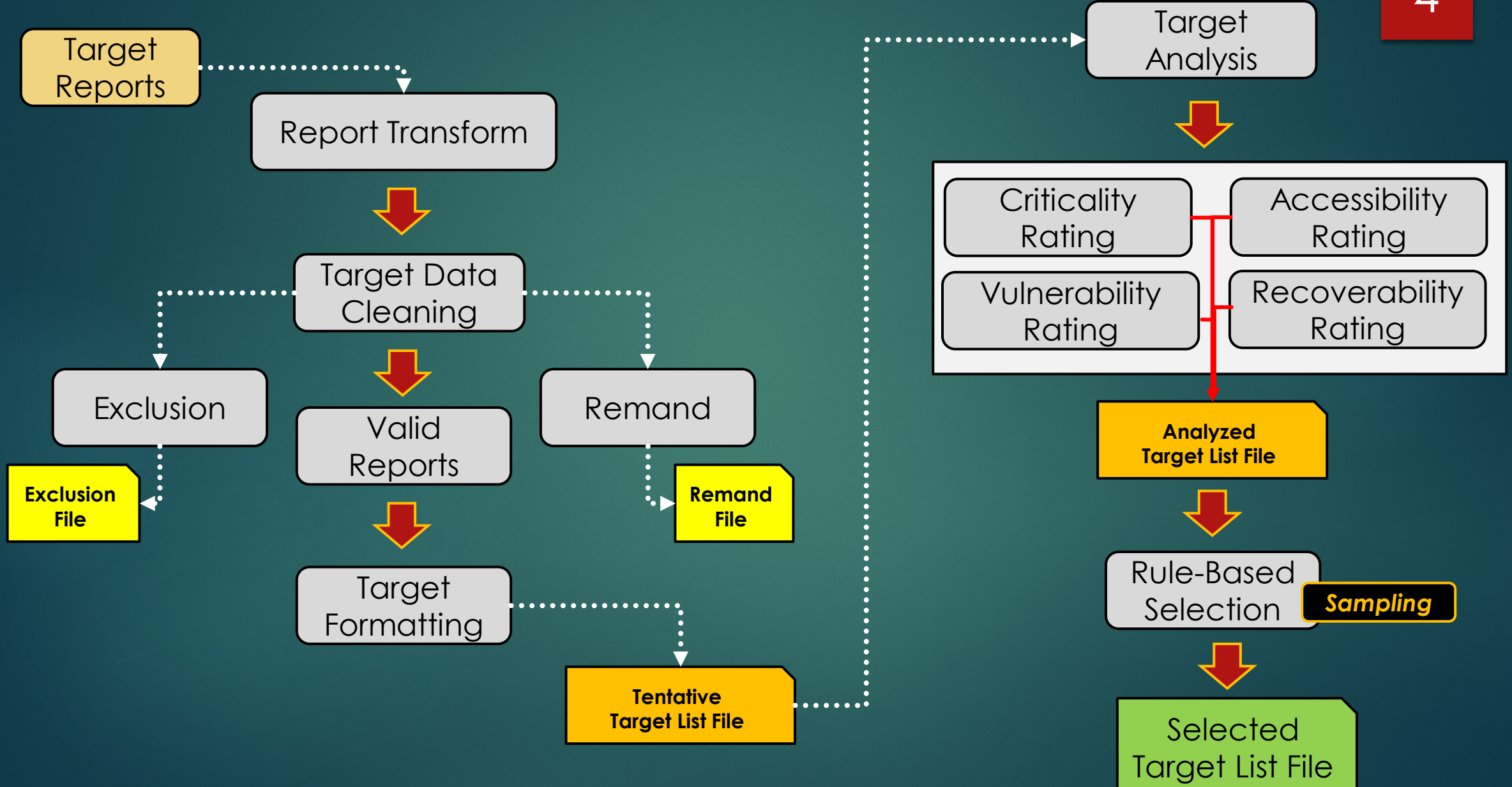
Problem Synopsis

3

- ▶ For our engineers to refine the ability of our Targeting AI system, we need to build a model of the Targeting Environment and provide a foundation for its analysis.
- ▶ To perform this analysis:
 - ▶ STEP 1: Shape Raw Data
 - ▶ STEP 2: Build a Data Model for Raw Data
 - ▶ STEP 3: Create an Analysis of the Data Model
 - ▶ STEP 4: Provide the hyperparameters for our Targeting AI system by examining the analysis of the model.

Targeting Model

4



Targeting Process

5

- **PART I Data Cleaning and Shaping**

- **STEP 1:** Target Reports are read into the Target Format
- **STEP 2:** Target that have prohibited attribute values are excluded from the possible target list file and recorded in the Exclusion file.
- **STEP 3:** Target that have missing attribute values are excluded from the possible target list file and recorded in the Remand file.
- **STEP 4:** Remaining targets are 'Nominated' and assigned a Target Number.
- **STEP 5:** After target are formatted into a useable form and written to the Tentative Target List file.

- **PART II Data Modelling**

- **STEP 6:** Analyze the Tentative Target List File.
- **STEP 7:** Use Rule-Based Sampling to get a sample of the Targets
- **STEP 8:** Produce a final Selected Target List file.

Target Report

6



1. Target Report Number
2. Target Report Time
3. Target Reported By
4. Target Type
5. Target Mobility
6. Target Length
7. Target Width
8. Target Defense
9. Target Protection
10. Target Activity
11. Target Location X
12. Target Location Y

Target Report File

ATargets

1	129561#74	6802	Ground	Zeta	Fixed	53	75	High	Fortified	Semi-Active	638	981
2	863451#93	6746	Signal	Delta	Mobile	55	92	Medium	Bastioned	Semi-Active	643	407
3	647593#95	8641	Ground	Alpha	Static	92	74	None	Fortified	Semi-Active	930	671
4	138534#92	4235	Ground	Gamma	Static	65	51	Medium	Bastioned	Semi-Active	172	58
5	199236#55	1989	Signal	Epsilon	Static	59	96	None	Fortified	Dormant	326	21
6	847746#54	5857	Airborne	Gamma	Static	72	77	High	Reinforced	Semi-Active	157	17
7	985641#16	2351	Ground	Delta	Mobile	25	41	Medium	Reinforced	Active	884	723
8	545338#56	417	Signal	Delta	Static	68	89	Medium	Reinforced	Dormant	863	59
9	586141#11	8232	Ground	Alpha	Static	55	62	None	Barricaded	Semi-Active	605	476
10	159732#32	2015	Ground	Alpha	Fixed	70	61	Low	Bastioned	Active	415	745

Target Reports

7

Corrupted Target Report

```
825854#48 4709 Signal UNK Mobile 59 93 High UNK Dormant 922 783
385973#19 7463 Ground UNK Static 72 113 UNK Fortified Dormant 68 603
```

Remand
File

Corrupted Target Reports are recorded in the **Remand** file
Corrupted Target Reports contain one or more **UNK** values

Excluded Target Report

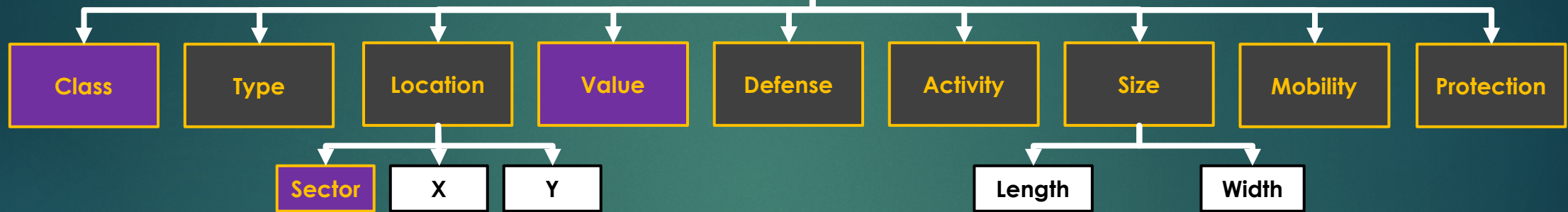
```
736962#59 4752 Ground Zeta Mobile 59 96 Trivial Bastioned Semi-Active 154 344
```

Exclusion
File

Excluded Target Reports are recorded in the **Excluded** file
Excluded Target Reports are ones which contain **Zeta** (Target Type) **Mobile** (Target Mobility)

Target Attributes

8



- **Class** ['Strategic', 'Operational', 'Tactical']
- **Type** ['Alpha', 'Beta', 'Gamma', 'Delta', 'Epsilon', 'Zeta', 'Eta', 'Theta', 'Iota ']
 - Alpha, Beta, Gamma = **Tactical Class**
 - Delta, Epsilon Zeta = **Operational Class**
 - Eta, Theta, Iota = **Strategic Class**
- **Location** [X = [0-1000] Y = [0-1000]]...see slide to determine Sector
- **Value** ...see slide for formula
- **Defense** = ['High', 'Medium', 'Low', 'None']
- **Activity** = ['Active', 'Semi-Active', 'Dormant']
- **Size** = [Length = [5-100] Width = [5-150]]
- **Mobility** = ['Static', 'Mobile', 'Fixed']
- **Protection** = ['Bastioned', 'Fortified', 'Reinforced', 'Barricaded']

Battle Map

9



Target Value

10

Class	
State	Value
Strategic	60
Operational	40
Tactical	20

Activity	
State	Value
Active	15
Semi-Active	10
Dormant	5

Mobility	
State	Value
Fixed	3
Static	2
Mobile	1

Protection	
State	Value
Bastioned	5
Fortified	4
Reinforced	3
Barricaded	2

Defense	
State	Value
High	5
Medium	4
Low	3
Trivial	1

$$V = \frac{Class \times Activity \times Mobility}{\sqrt{Protection \times Defense}}$$

11

Tentative Target List File

[illegible]

Target Analysis

12

$$\textit{Criticality} = \textit{Class} \times \textit{Protection} \times \textit{Defense}$$

$$\textit{Accessablity} = \textit{Range} \times \textit{Defense} \times \textit{Mobility}$$

$$\textit{Vulnerablity} = \textit{Accessability} \times \frac{1}{\textit{Protection}}$$

$$\textit{Recoverablity} = \textit{Criticality} \times \frac{1}{\textit{Vulnerability}}$$

Class	
State	Value
Strategic	60
Operational	40
Tactical	20

Protection	
State	Value
Bastioned	5
Fortified	4
Reinforced	3
Barricaded	2

Defense	
State	Value
High	5
Medium	4
Low	3
Trivial	1

Mobility	
State	Value
Fixed	3
Static	2
Mobile	1

Range is calculated from the 0,0 Coordinate

Target Analysis

13

```
TGT-1 400 7364.79 3682.39 0.11
TGT-2 240 1861.44 465.36 0.52
TGT-3 360 1763.16 587.72 0.61
TGT-4 320 12233.7 3058.43 0.1
TGT-5 1500 19598.46 3919.69 0.38
```

Analyzed
Target List File



Target Number	Criticality	Aceessibility	Vulnerability	Recoverability

Rule Based Selection

14

Rule	Criteria	Maximum Count
1	Criticality is within 1 Sigma of the mean of Criticality	25% of Targets that meet criteria
2	Accessibility is above the mean of Accessibility	75% of Targets that meet criteria
3	Strategic Target in the Q3 of Range	100% of Targets that meet criteria
4	Recoverability is below the mean of Recoverability	50% of Targets that meet criteria
5	Vulnerability is above the mean	25% of Targets that meet criteria

Do not include duplicate Targets in the Selected Target List file



Selected
Target List File

Model Analysis

15

1. *What is the probability that any target located in Sector A is Strategic.
 1. Use Tentative Target File
2. *What is the probability that any target located in Sector C is of class 'Strategic' given that its Mobility is 'Static'.
 1. Use Tentative Target File
3. *Calculate the Posterior for Strategic Targets.
 1. Use Tentative Target File
4. What is the Euclidean Distance between the Mean Feature Vector and every Target in the Selected Target List file.
 1. Use Analyzed Target File

***Extra Credit for Generic Solutions**