

## Workshop 8: Week of October 28

### Week 8 Workshop Response Question

(submit on Blackboard by 11:59 PM Sunday, 11/3) (3 points)

Discuss the Industrial Safety Assessment case on the following pages. Choose one of the Managers from the final table (Kim, Lee, Jones, Garcia) and create a brief set of recommendations (3 or 4 bullet points) for that manager's safety priorities and interventions, based on the facility risk profiles in their divisions.

Please submit your response (only one per team is required) on Blackboard as a .doc or .pdf document.

This exercise involves using PCA to identify meaningful patterns in complex datasets and applying K-means clustering for segmentation. You will use these techniques again in Case 3 with NLS data. Along with completing this exercise, use the Workshop meeting to develop an analysis plan and timeline for your Case 3 work. I highly recommend that your team aim to complete the analytics by the next (Week 9) Workshop, so you can dedicate that time next week to discussing interpretation, visualization, and client presentation strategies.

#### Case setting:

Precision Engineering Solutions operates multiple manufacturing facilities across a region. Following recent safety incidents and pressure from regulatory bodies, Precision has initiated a comprehensive safety assessment program. This program aims to evaluate the risk profile of each facility, summarize risk levels by division and manager, and prioritize resources to improve workplace safety across the organization.

An analytics team has been engaged to assess operational, safety, and environmental data across 120 facilities. Using Principal Component Analysis (PCA) and K-means clustering, the team will segment the facilities into groups with unique risk profiles. Precision will use these insights to customize safety interventions for each facility segment and evaluate manager profiles. Potential recommended interventions include on-site training, equipment upgrades, and partnerships with local emergency responders.

The data files contain 6 variables for the 120 Precision facilities. The columns in the data file are described in the table below.

## Data Dictionary

Field	Description
Facility_ID	Unique identifier for each facility.
Employee_Density	Number of employees per square foot, representing the facility's crowding level.
Avg_Equipment_Age	Average age of major equipment in years, with older equipment posing higher risks.
Safety_Violations_Last_5_Years	Number of recorded safety violations or incidents per 100 employees over the past five years.
Dist_to_Nearest_Emergency_Service	Distance in miles to the nearest emergency service, as proximity can impact response time.
Total_Annual_Production	Total production output in units, indicating operational scale.
Maintenance_Expenditure_Per_SqFt	Amount spent on maintenance per square foot, a measure of upkeep and safety investment.

## A few rows of the data file

Facility_ID	Employee Density	Avg Equipment Age	Safety Violations Last 5 Years	Dist to Nearest Emergency Service	Total Annual Production	Maintenance Expenditure Per SqFt
FAC_1	80	7	2	2.5	5000	15
FAC_2	120	3	5	1	8000	20
FAC_3	90	10	8	3.5	4500	12
FAC_4	75	8	3	2	5500	14
FAC_5	110	5	1	0.5	7000	18
FAC_6	95	6	4	1.5	6000	16
FAC_7	130	12	6	3	9000	11
FAC_8	85	4	3	1	4800	17
FAC_9	105	9	7	2.5	7500	13
FAC_10	115	7	2	1.5	6200	19

PC1: general information of equipment--upgrades  
 PC2: employee density, production--operational risks--over production  
 PC3: violations--safety

### Principal Component Output:

Variable	PC1	PC2	PC3	PC4	PC5	PC6
Employee_Density	0.03	0.69	-0.02	0.23	-0.34	0.12
Avg_Equipment_Age	0.51	0.03	-0.46	-0.12	0.28	-0.42
Safety_Violations_Last_5_Years	0.43	0.12	0.84	0.45	-0.16	0.25
Dist_to_Nearest_Emergency_Service	0.52	-0.14	0.00	-0.36	0.39	0.37
Total_Annual_Production	0.07	0.68	-0.15	0.22	-0.15	-0.29
Maintenance_Expenditure_Per_SqFt	-0.52	0.13	0.23	-0.4	0.31	0.33
Explained Variance Ratio	36.8%	19.7%	18.8%	8.8%	9.2%	6.6%

### K-Means Output (first 3 Principal Components)

Cluster	PC1	PC2	PC3	Size
0	-0.55067	-1.21118	-0.14136	0.523
1	2.507734	0.498862	0.185481	0.189
2	-1.7735	1.116045	0.002999	0.288

### Mix of Clusters by Precision Division and Manager

Division	Manager	Cluster 0	Cluster 1	Cluster 2
North East	Kim	80%	8%	12%
North East	Jones	65%	10%	25%
South West	Lee	40%	25%	35%
South West	Garcia	55%	15%	30%
Midwest	Kim	30%	20%	50%
Midwest	Jones	50%	20%	30%
Pacific	Lee	70%	10%	20%
Pacific	Garcia	20%	15%	65%