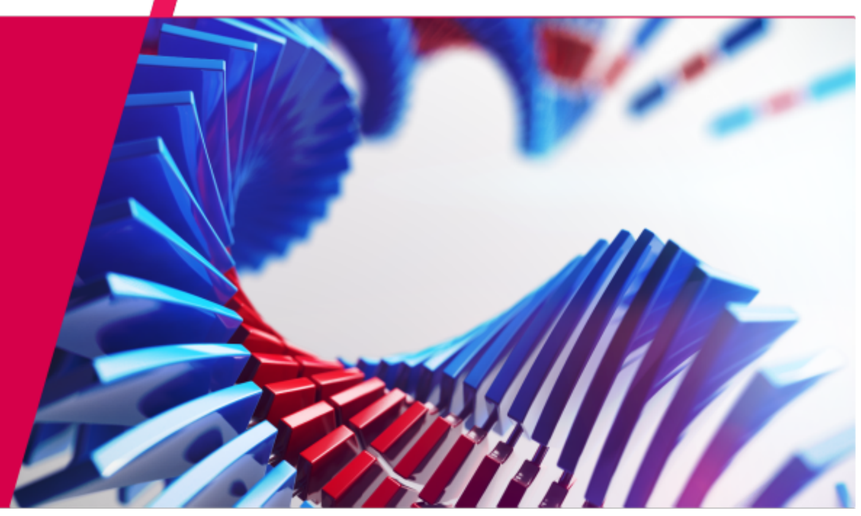
2DMT00 Applied statistics

Technische Universiteit Eindhoven University of Technology

Research Report – Group 11



Mingpeiyu Zhang - 1018903
Niels Hellinga - 0977787
Vishal Chouksey - 1034346

Problem Statement

An accident occurred in the factory and there is concern of release of a toxic substance at different site of the factory. Hence the levels of PPM were measured using 2 sensors on 7 sites hourly, up to 4 hours

Questions

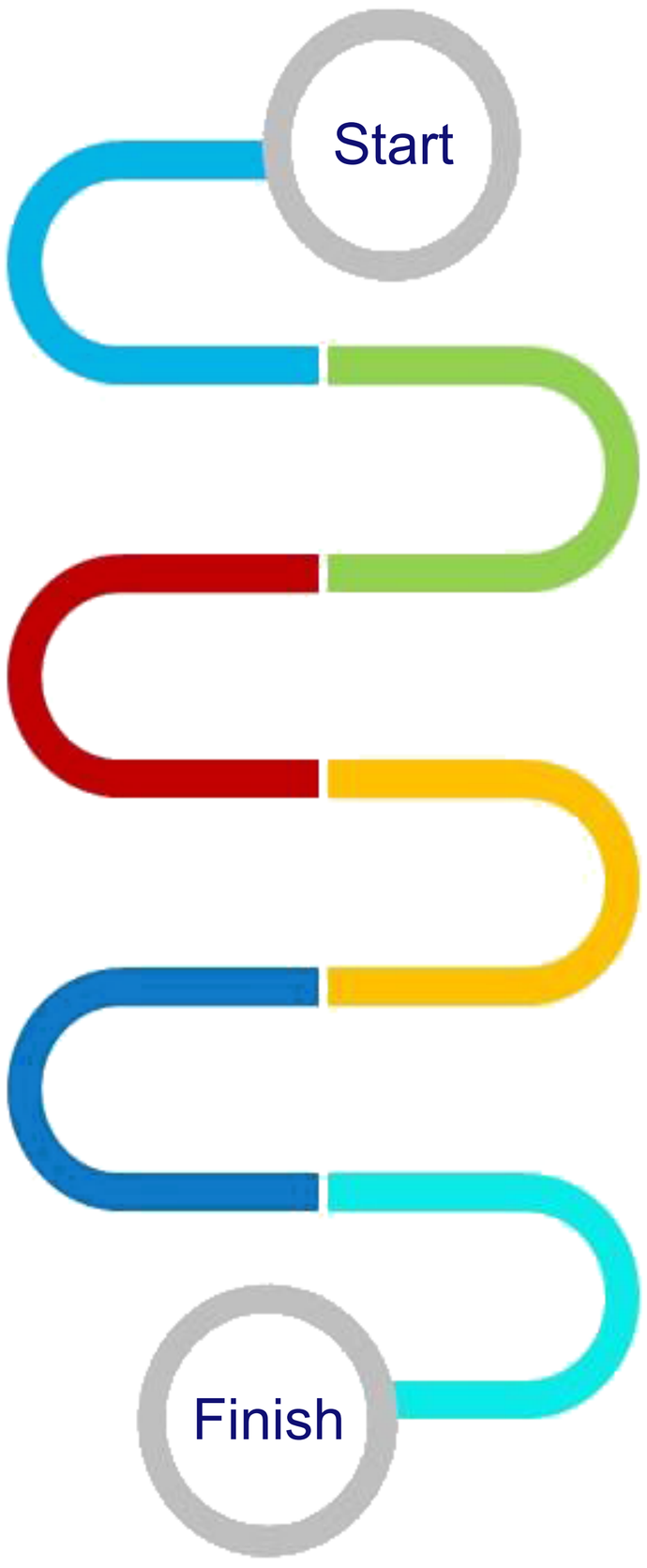
- Is/are there outlier/s in the data set?
- Are the data normally distributed?
- Do the PPM levels of the sites have different locations and variances compared to the PPM levels before the accident?

Analysis & Procedure

01

Summary Statistics:

Site	n	min	max	mean	median	sd	var
All	3378	0,000	25,681	3,748	2,241	4,063	16,506
Before	120	0,036	10,164	2,132	1,420	2,128	4,526
Site1	480	0,000	10,045	2,786	2,098	1,989	3,958
Site2	480	0,000	4,482	1,154	1,019	0,873	0,761
Site3	459	0,140	19,567	5,790	3,691	5,079	25,799
Site4	459	0,000	25,681	7,917	7,076	6,265	39,250
Site5	480	0,496	8,452	2,619	2,234	1,486	2,209
Site6	480	0,000	16,482	3,019	1,940	3,286	10,801
Site7	420	0,289	12,873	3,608	3,105	2,272	5,164



02

Paired sample test:
Check if Sensor 1 and Sensor 2 have the same distribution. Sensors of Site Before, Site 1 and Site 3 have the same distribution, while the distribution of the two sensors in the other 7 sites are significantly different.

03

Outliers Test:
We choose only to take the the testing result of Doornbos test. Since all the other tests assume that the samples come from a normally distributed population which we do not know so far. The results of other tests can not be trusted. We then remove the 2 detected outliers

04

Normality Test:
We have two cases where we can detect normality, namely Site Before with the transformed LOGPPM values and Site 5 for Sensor 1 with the 'normal' PPM values. Since the rest is not normally distributed and we cannot compare Site 5 to Site Before, we will need to use non-parametric tests.

05

Variance & Location:
Site 1, Site 5 (Sensor 2), Site 6 and Site 7 do not have different variances compared to the PPM levels before the accident. The location has changed for all the sites when compared to the Site Before.

06

Results:
Site 1 and 2 are not worse than normal level (safe).
Site 3, 4 and 6 are worse than normal level (unsafe).
Site 5 and 7 We cannot say.

Site 1 **Site 2** **Site 3** **Site 4** **Site 5** **Site 6** **Site 7**