

## **GEM Trailblazer Summer Programme**

### **DETAILED COURSE INFORMATION**

<b>Academic Year</b>	:	AY2022-2023
<b>Session</b>	:	Summer 2023
<b>Course Code &amp; Title</b>	:	CS2400 Foundations of Information Analytics
<b>Academic Units</b>	:	3 AUs
<b>Pre-requisite (if any)</b>	:	NIL
<b>Delivery mode</b>	:	Physical
<b>Taught by</b>	:	Dr Lee Chu Keong

#### **Brief Course Description**

Today, many organisations generate, and collect (“harvest”), unimaginable quantities of data of all types. However, merely collecting lots of data is pointless. The critical step is to analyse the data so that it can be transformed into information and action. The key idea is to transform data in such a way that it can be used for business advantage. An important tool that enables this transformation is statistics. This is the subject matter of this course. Statistics will be presented in a mathematically friendly and non-threatening manner. The course emphasises conceptual understanding of the material, and not on the exact keystrokes needed to accomplish specific statistical tests.

#### **1. Intended Learning Objectives (ILOS)**

- To sensitise you to the fact that data is all around us
- To start you thinking about the opportunities for transforming data into action (new products / processes / ...)
- To lay the statistical foundations for data analytics so that you can transform the data into actionable information

## 2. Course Syllabus/Topics

Lesson	Topic/s
1	Introduction
2	Descriptive Statistics
3	Exploratory Data Analysis
4	Inferential Statistics: Parametric Methods I
5	Inferential Statistics: Parametric Methods II
6	Mid-semester Test
7	Correlation and Regression
8	Inferential Statistics: Non-parametric Methods I
9	Inferential Statistics: Non-parametric Methods II
10	Final Test

## 3. Learning Outcomes

- To understand how data analytics is used in the real world
- To be able to describe data using measures of central tendency, spread and shape
- To be able to read and produce visualisations for data
  - traditional visualisation methods
  - non-traditional visualisation methods
- To be able to perform parametric statistical tests
- To be able to perform non-parametric statistical tests
- To be able to perform correlation analysis for data in the form of values and ranks
- To be able to perform linear regression, multiple regression and polynomial regression

## 4. Course Assessments

Components	Group/Individual	Weighting
Mid-Semester Test	Individual	20%
Group Assignment	Group	20%

Individual Assignment	Individual	20%
Final Test	Individual	40%
<b>Total</b>		<b>100%</b>

## 5. References/Recommended Reading List

### Scholarly Articles

- Brasseur, L. (2009). Florence Nightingale's Visual Rhetoric in the Rose Diagrams. *Technical Communication Quarterly*, 14(2), 161–182.
- Crowther, T.W., Glick, H.B., Covey, K.R. et al. (2015). Mapping tree density at a global scale. *Nature*, 525, 201-205.
- Donoho, D. (2017). 50 Years of Data Science. *Journal of Computational and Graphical Statistics*, 26(4), 745-766.
- Ferguson, D. (2013). How supermarkets get your data – and what they do with it. *The Guardian*.
- Gurin, J. (2014). Opening Business Innovation with Open Data. *Business Horizon*, 12, 42–49.
- Hamilton, A.C., Donnelly, D.W., Fitzpatrick, D., & Coleman, H.G. (2022). Early-Onset Cancers in Adults: A Review of Epidemiology, Supportive Care Needs and Future Research Priorities. *Cancers (Basel)*, 14(16), 4021.
- Hess, A. (May 14, 2017). Open Secrets. *The New York Times Magazine*. New York: New York Times.
- Khan, M.A., Uddin, M.F., & Gupta, N. (2014). Seven V's of Big Data: Understanding Big Data to Extract Value. *Conference of the American Society for Engineering Education*.
- Loukides, M. (2010). *What is Data Science?* Sebastopol, CA: O'Reilly.

### Books

- Black, E. (2012). *IBM and the Holocaust*. Dialog Press.
- Johnson, S. (2006). *The Ghost Map: The Story of London's Most Terrifying Epidemic and How It Changed Science, Cities, and the Modern World*. Riverhead Books.
- Winchester, S. (2011). *The Professor and the Madman: A Tale of Murder, Insanity, and the Making of the Oxford English Dictionary*. HarperCollins.

## Videos

*The Best Stats You've Ever Seen (A TED Talk by Hans Rosling)*  
(<https://www.youtube.com/watch?v=hVimVzgtD6w>)

*The Joy of Stats. By Hans Rosling* (<http://www.gapminder.org/videos/the-joy-of-stats/>)

*Trash Trail: Episode 5 (Data: Finale)* (<https://video.toggle.sg/en/video/series/trash-trail/ep5/478356>)

## Statistical Tables

Lindley, D.V., & Scott, W.F. (1995). *New Cambridge Statistical Tables (2nd ed.)*. Cambridge University Press.

## 6. Other requirements

Nil

## 7. Instructor Details

Course Instructor	Office	Email
Dr Lee Chu Keong		ascklee@ntu.edu.sg