|  | | **COURSE** | **STATISTICS FOR DATA SCIENCE** | | --- | --- | | **TUTOR** | **Wycliffe Bosire** | | **YEAR** | **2023** | |
| --- | --- | --- | --- | --- | --- | --- | --- |

| Concept map of unit | | | | | |
| --- | --- | --- | --- | --- | --- |
| Why Statistics is important for Data Science | | | | | |
| Key learning(s) | | unit essential questions | | optional instructional tools | |
| Statistics for Data Science | | Where do we apply statistical techniques in Data Science and why? | | Online resources available | |
| Concept #1 | Concept #2 | | Concept #3 | | Concept #4 |
| **Descriptive Statistics** | **Probability Theory** | | **Inferential Statistics** | | **Data Visualization** |
| Lesson essential questions #1 | Lesson essential questions #2 | | Lesson essential questions #3 | | Lesson essential questions #4 |
| * Measures of Central Tendency * Measures of Dispersion | * Probability in Statistics * Common probability distributions | | * Hypothesis testing * Significance levels and p-values * Characteristics of good features/estimators | | * Introduction to Matplotlib, Seaborn and Power BI * Choosing the right visualizations |
| vocabulary #1 | vocabulary #2 | | vocabulary #3 | | vocabulary #4 |
| * Mean, Median, Mode * Range, variance, standard deviation | * Discrete and continuous distributions | | * Hypothesis testing * ANOVA, Estimators/ Features * Predictive Statistics | | * Charts and graphs |
|  | | | | | |
|  | | | | | |