# KAUTILYA SCHOOL OF PUBLIC POLICY
GITAM (Deemed to be University)
Rudraram, Patancheru Mandal
Hyderabad, Telangana 502329

| | | |
|---|---|---|
| **Course Code**: PPOL6731 | **Course Title**: Statistics for Public Policy | |
| **Trimester**: 1 | **Course Type**: Core | **Credits**: 3 |
| **Home Program(s)**:  MPP | **Batch/Academic Year**: 2023-2025 | |
| **Course Lead:**  Dr. Vishnu S Pillai | **Assigned T/RA:** | |

## Course Description

Statistical thinking, comprehension, and competence are more crucial than ever in today's increasingly complex environment. Thus, understanding the mathematical and statistical concepts is necessary to understand the various policy frameworks. This course introduces students to statistical techniques used in Policy analysis and evaluation. It provides a solid understanding of the ideas and criteria for performing quantitative research. The emphasis will be on establishing a conceptual grasp of statistics and data analysis abilities and learning how to apply them to policy analysts' challenges.

This course will focus on Python as the programming language for data analysis. No prerequisite programming knowledge is required for this course, and students will receive valuable hands-on experience using Python for data analysis. The course content lays the groundwork for students to pursue more advanced topics in the discipline as per their career and research requirements. The techniques discussed in this course will significantly help students learn and understand theories and applications discussed in advanced courses on data analysis in the Public Policy domain.

## Learning Objectives

1. Engage students in sound statistical thinking and reasoning.
2. Explore multiple representations of topics, including graphical, symbolic, numerical, oral, and written.
3. Provide the students with the skills required to be proficient in statistics.
4. Provide the students with basic programming skills for data analysis using Python.

## Course Outcomes
On successful completion of this course:

1. Students will be equipped with the skills in describing, analysing, and interpreting policy data using statistical methods.

2. Develop research questions and test hypotheses for a research project.
3. Understand practical applications of statistical concepts that can be used in the public policy sector.
4. They can relate and apply statistical methodology to policy analysis using Python.
5. The students can produce policy papers and reports using statistical analysis.

## Course Schedule

| Unit 1 | Sessions: 2 | Laying the foundation |
|---|---|---|

1. Why Statistics for Public Policy?
2. Descriptive and inferential statistics.
3. Types of variables and scales of measurement.
4. Descriptive statistics - measures of central tendencies, measures of spread.
5. Basics of Probability.

| Unit 2 | Sessions: 4 | Introducing Python |
|---|---|---|

1. Introducing Python.
2. Python – conditions, loops, and functions.
3. Python - univariate descriptive statistics - measures of central tendency and spread.
4. Python - NumPy package - creating and working with n-dimensional arrays - basics, filtering, slicing, and other array operations.

| Unit 3 | Sessions: 4 | Understanding probability distributions |
|---|---|---|

1. Probability distributions.
2. Python NumPy package – simulations and analysing distributions.

| Unit 4 | Sessions: 2 | Data collection |
|---|---|---|

1. Methods of data collection – introducing experimental, quasi-experimental, and non-experimental methods.
2. Various sampling methods, errors, and biases.

| Unit 5 | Sessions: 4 | Entering the world of data analysis with Python |
|---|---|---|

| | | |
|---|---|---|
| 1. Python Pandas introduction - data structures and basic functionalities. 2. Data loading in Python. 3. Data cleaning and dealing with missing data. 4. Visualising data in Python (Matplotlib, Seaborn). | | |
| **Unit 6** | **Sessions: 2** | **Hypothesis testing** |
| 1. Hypothesis testing - basic steps. 2. Hypothesis testing of means when the population standard deviation is known/unknown. 3. Analysis of Variance – ANOVA. | | |
| **Unit 7** | **Sessions: 6** | **Prediction** |
| 1. Setting the context for linear regression. 2. Linear regression model using Python. 3. Multiple linear regression. 4. Introducing the basics of Machine Learning. | | |

As of Oc 2024