



College of Engineering & Applied Sciences

# CSPB 3022

*Introduction To Data Science With Probability And Statistics*

*Class Notes*

UNIVERSITY OF COLORADO

2024

Sections			
Data Science Lifecycle, Python And Pandas .	2	Probability: Independence, Simulation,	
1.0.1 Piazza . . . . .	2	Random Variables . . . . .	7
1.0.2 Lectures . . . . .	2	5.0.1 Piazza . . . . .	7
1.0.3 Assignments . . . . .	3	5.0.2 Lectures . . . . .	7
1.0.4 Quiz . . . . .	3	5.0.3 Assignments . . . . .	7
Pandas, Exploring And Cleaning Tabular		5.0.4 Quiz . . . . .	7
Data . . . . .	4	Exam 1 . . . . .	8
2.0.1 Piazza . . . . .	4	6.0.1 Piazza . . . . .	8
2.0.2 Lectures . . . . .	4	6.0.2 Lectures . . . . .	8
2.0.3 Assignments . . . . .	4	6.0.3 Quiz . . . . .	8
2.0.4 Quiz . . . . .	4	Expected Value, Variance - Discrete And	
Exploratory Data Analysis And Visualization		Continuous RV . . . . .	9
5		7.0.1 Piazza . . . . .	9
3.0.1 Piazza . . . . .	5	7.0.2 Lectures . . . . .	9
3.0.2 Lectures . . . . .	5	7.0.3 Assignments . . . . .	9
3.0.3 Assignments . . . . .	5	Covariance And Correlation, Independence	
3.0.4 Quiz . . . . .	5	- Sampling . . . . .	10
Visualization And Introduction To Probability		8.0.1 Piazza . . . . .	10
6		8.0.2 Lectures . . . . .	10
4.0.1 Piazza . . . . .	6	8.0.3 Assignments . . . . .	10
4.0.2 Lectures . . . . .	6	8.0.4 Quiz . . . . .	10
4.0.3 Assignments . . . . .	6	CLT And Hypothesis Testing . . . . .	11
4.0.4 Quiz . . . . .	6	9.0.1 Piazza . . . . .	11
		9.0.2 Lectures . . . . .	11
		9.0.3 Assignments . . . . .	11

## Week 1: (1/15 - 1/21)

# Data Science Lifecycle, Python And Pandas

### 1.0.1 Piazza

Must post / respond to two Piazza posts. (1/23/24)

### 1.0.2 Lectures

The lecture videos for this week are:

- Course Overview  $\approx$  54 min.
- Pandas Part 1  $\approx$  54 min.
- Pandas Part 2  $\approx$  54 min.
- Python Walkthrough  $\approx$  47 min.
- Learning to Use Jupyter Notebooks  $\approx$  9 min.
- Using Markdown In Jupyter Notebooks  $\approx$  11 min.
- Writing Mathematics In Jupyter Notebooks  $\approx$  16 min.
- Beginning Python  $\approx$  44 min.
- Python - Basic NumPy  $\approx$  18 min.
- Quick Calculus Refresher  $\approx$  21 min.

The lecture notes for this week are:

- [Course Overview Data Science Lifecycle](#)
- [Pandas Part 1](#)
- [Pandas Part 2](#)
- [L<sup>A</sup>T<sub>E</sub>X](#)

### 1.0.3 Assignments

The assignment for this week is:

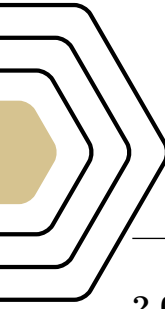
- The assignment for this week is **Data Science Lifecycle, Python And Pandas**. [\(1/23/24\)](#)

### 1.0.4 Quiz

The quizzes for this week are:

- [Math Concept Quiz](#) • [Finalized Math Concept Quiz](#) [\(1/30/24\)](#)





# Pandas, Exploring And Cleaning Tabular Data

---

## 2.0.1 Piazza

Must post / respond to two Piazza posts. (1/30/24)

## 2.0.2 Lectures

The lecture videos for this week are:

- Pandas Part 3  $\approx$  54 min.
- Pandas Part 4  $\approx$  54 min.
- Discussion 2 Video Walkthrough: Basics with Pandas and In-Depth with NumPy  $\approx$  73 min.

The lecture notes for this week are:

- Pandas Part 3
- Pandas Part 4
- Permutations And Combinations
- Section 7.1 Intro To Discrete Probability
- Section 7.2 Probability Theory And Applications

## 2.0.3 Assignments

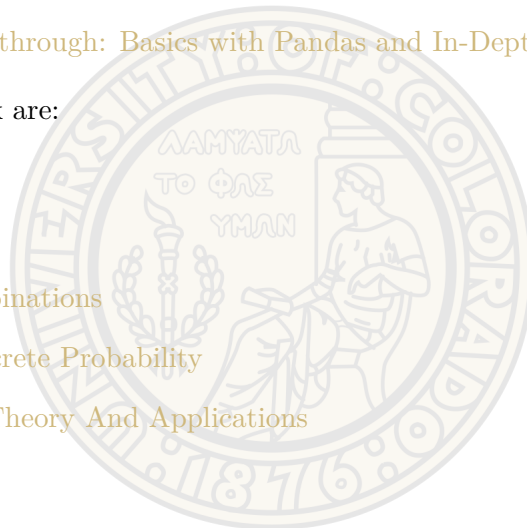
The assignment for this week is:

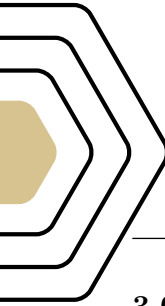
- The assignment for this week is **Pandas, Exploring And Cleaning Tabular Data**. (1/30/24)

## 2.0.4 Quiz

The quizzes for this week are:

- Python And Pandas Quiz • Finalized Python And Pandas Quiz (1/30/24)





# Exploratory Data Analysis And Visualization

---

## 3.0.1 Piazza

Must post / respond to two Piazza posts. (2/6/24)

## 3.0.2 Lectures

The lecture videos for this week are:

- Lecture 6: Data Wrangling And EDA  $\approx$  54 min.
- Lecture 7: EDA And Visualization  $\approx$  54 min.
- Lecture 8: Quiz Walkthrough And Visualization  $\approx$  54 min.
- Discussion 3: More With Pandas  $\approx$  50 min.

The lecture notes for this week are:

- Pandas Part 4
- Exploratory Data Analysis
- Visualization

## 3.0.3 Assignments

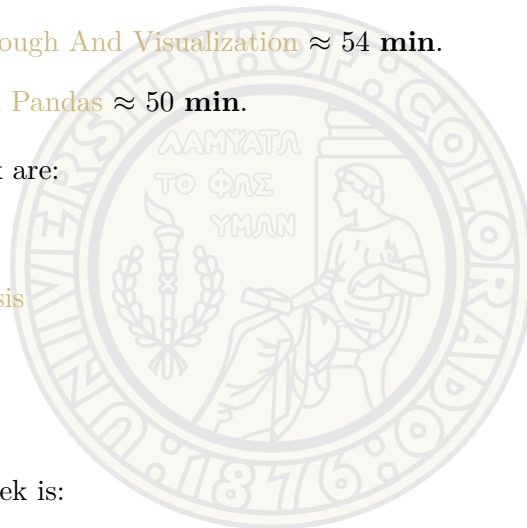
The assignment for this week is:

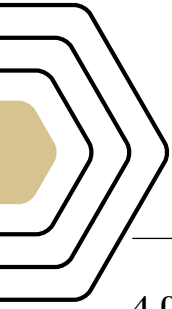
- The assignment for this week is **Exploratory Data Analysis And Visualization**. (2/6/24)

## 3.0.4 Quiz

The quizzes for this week are:

- Math And Pandas Quiz • Finalized Math And Pandas Quiz (2/6/24)





# Visualization And Introduction To Probability

---

## 4.0.1 Piazza

Must post / respond to two Piazza posts. (2/13/24)

## 4.0.2 Lectures

The lecture videos for this week are:

- Visualization and Introduction to Probability  $\approx$  54 min.
- Probability Part 1  $\approx$  54 min.
- Quiz Walkthrough; Probability  $\approx$  54 min.
- Bike Sharing EDA and Visualization  $\approx$  49 min.

The lecture notes for this week are:

- Intro To Probability Lecture Notes
- Probability II - Total Probability, Bayes Rule And Independence Lecture Notes

## 4.0.3 Assignments

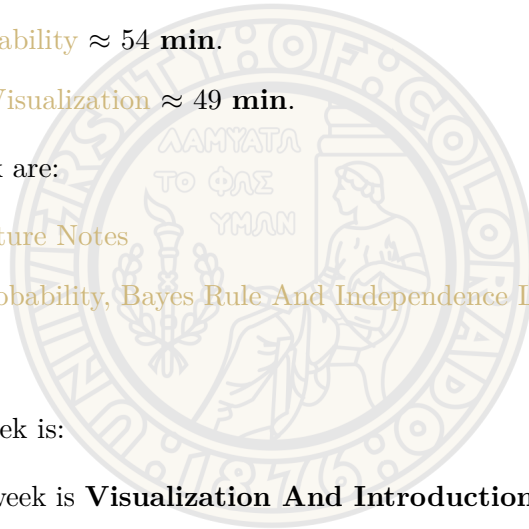
The assignment for this week is:


- The assignment for this week is **Visualization And Introduction To Probability**. (2/13/24)

## 4.0.4 Quiz

The quizzes for this week are:

- Quiz 4 • Finalized Quiz 4 (2/13/24)





# Probability: Independence, Simulation, Random Variables

---

### 5.0.1 Piazza

Must post / respond to two Piazza posts. (2/20/24)

### 5.0.2 Lectures

The lecture videos for this week are:

- Probability Cont. Independence And Simulation  $\approx 54$  min.
- Probability Cont. Independence And Simulation Jupyter Notebook Examples  $\approx 27$  min.
- Random Simulation Example  $\approx 45$  min.
- Independence Cont. Discrete Random Variables  $\approx 54$  min.
- Quiz Walkthrough, Common Discrete RVs, Expectation  $\approx 54$  min.

The lecture notes for this week are:

- Independence Lecture Notes
- Discrete Random Variables Lecture Notes
- Expected Value And Variance Lecture Notes

### 5.0.3 Assignments

The assignment for this week is:

- The assignment for this week is **Probability: Independence, Simulation, Random Variables.** (2/20/24)

### 5.0.4 Quiz

The quizzes for this week are:

Quiz 5 • Finalized Quiz 5 (2/20/24)

---

## Week 6: (2/19 - 2/25)



### Exam 1

---

#### 6.0.1 Piazza

Must post / respond to two Piazza posts. (2/27/24)

#### 6.0.2 Lectures

The lecture videos for this week are:

- Expected Value, Variance, Discrete And Continuous RV  $\approx 54$  min.

The lecture notes for this week are:

- Expected Value, Variance Discrete And Continuous RV Lecture Notes

#### 6.0.3 Quiz

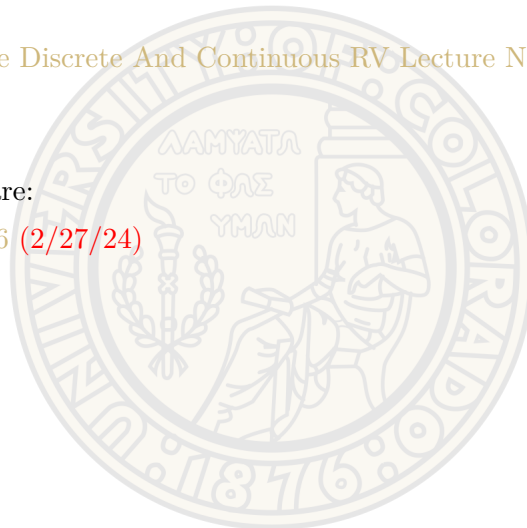
The quizzes for this week are:

- Quiz 6 • Finalized Quiz 6 (2/27/24)

#### Exam

The exam for this week is:

- Exam 1 (2/23/24)







## Expected Value, Variance - Discrete And Continuous RV

---

### 7.0.1 Piazza

Must post / respond to two Piazza posts. (3/5/24)

### 7.0.2 Lectures

The lecture videos for this week are:

- Random Variables And Distributions Notebook Walkthrough  $\approx 83$  min.
- RV And Distributions, Expectation And Variance  $\approx 54$  min.
- Joint Distributions, Covariance, Independence  $\approx 54$  min.
- Covariance And Correlation, Independence  $\approx 54$  min.

The lecture notes for this week are:

- Expected Value, Variance Discrete And Continuous RV Lecture Notes
- More With RV - Discrete Vs. Continuous - Expected Value And Variance Lecture Notes
- Joint Distributions - Covariance And Correlation Lecture Notes
- Sampling Lecture Notes

### 7.0.3 Assignments

The assignment for this week is:

- The assignment for this week is **Expected Value, Variance - Discrete And Continuous RV**. (3/5/24)



## Covariance And Correlation, Independence - Sampling

---

### 8.0.1 Piazza

Must post / respond to two Piazza posts. (3/12/24)

### 8.0.2 Lectures

The lecture videos for this week are:

- Sampling  $\approx$  54 min.
- Central Limit Theorem  $\approx$  54 min.
- Hypothesis Testing  $\approx$  54 min.

The lecture notes for this week are:

- Sampling Lecture Notes
- Sample Statistics and The Central Limit Theorem Lecture Notes
- Hypothesis Testing Lecture Notes

### 8.0.3 Assignments

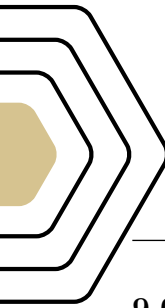
The assignment for this week is:

- The assignment for this week is **Covariance And Correlation, Independence - Sampling**. (3/12/24)

### 8.0.4 Quiz

The quizzes for this week are:

Quiz 7 • Finalized Quiz 7 (3/12/24)



# CLT And Hypothesis Testing

---

## 9.0.1 Piazza

Must post / respond to two Piazza posts. (3/19/24)

## 9.0.2 Lectures

The lecture videos for this week are:

- Hypothesis Testing  $\approx 54$  min.
- Testing Hypothesis And The Central Limit Theorem Walkthrough  $\approx 63$  min.
- A/B Testing  $\approx 54$  min.
- A/B Testing Walkthrough  $\approx 33$  min.
- Errors In Hypothesis Testing  $\approx 54$  min.

The lecture notes for this week are:

- Hypothesis Testing Lecture Notes
- AB Testing, Randomized Control Tests And Causality Lecture Notes
- Errors in Hypothesis Testing, PHacking Lecture Notes

## 9.0.3 Assignments

The assignment for this week is:

- The assignment for this week is **CLT And Hypothesis Testing**. (3/19/24)

