



שם הקורס

מבוא לסטטיסטיקה וניתוח נתונים עם R

מרצה

ד"ר עדי שריד

סמסטר

ב'

דרישות הקורס

מבוא להסתברות, שיטות מתמטיות 1

הרכב הציון הסופי

שיעורי בית חובה (ללא ציון), 40% פרויקט ניתוח נתונים, 60% מבחן

מבנה הקורס

תאריך / מס' שיעור	נושא השיעור ותכני השיעור (מטלות, רשימת קריאה, משימות וכיו"ב)
1	הקדמה: כיצד מתבצע מחקר סטטיסטי? שלבי העבודה מהתכנון, דרך ביצוע, והצגת הנתונים למקבלי החלטות. שימוש ב-R ככלי לניתוח נתונים.
2	אמידה נקודתית: שיטת אמד נראות מירבית, שיטת המומנטים. אמדים מוכרים ממוצע, סטיית תקן
3	המשך אמידה נקודתית / מעבדה ב-R.
4	רווח בר סמך לתוחלת, רווח בר סמך לפרופורציה.
5	מבחני השערות: לתוחלת, מדגם יחיד. מבחן z, מבחן t. סוגי טעויות.
6	מבחני השערות: שני מדגמים, מזווג/לא מזווג, פרופורציה ושונות.
7	מבחני השערות: טיב התאמה, אי-תלות.
8	המשך מבחני השערות / מעבדה ב-R.
9	רגרסיה לינארית פשוטה. מוטיבציה, דוגמאות, מודל.
10	רגרסיה לינארית פשוטה: המשך. מבחני השערות על המקדמים, מבחני השערות על הרגרסיה. משמעות, קורלציה.
11	רגרסיה לינארית מרובת משתנים: הרחבת המודל של רגרסיה לינארית פשוטה. מבחן F לרגרסיה, מולטיקולינאריות, דוגמאות.
12	רגרסיה לינארית מרובת משתנים: שימוש ב-Stepwise selection, אינטרוולים.
13	מעבדה ב-R / נושאים מתקדמים (ככל שיותר הזמן)

קריאת חובה

קריאת רשות

OpenIntro statistics is an introduction to statistics with R, it doesn't contain everything we will learn, but provides a good intro to some topics. Downloadable for free [here](#) (click on the "download sample" and the entire book downloads as a pdf file).

- Diez, D. M., Barr, C. D., & Cetinkaya-Rundel, M. (2012). OpenIntro statistics (pp. 174-175). OpenIntro.

R4DS (R for Data Science) is a highly recommended book for learning R, and specifically tidyverse which is a collection of useful packages for data science. The book is mostly "technical", i.e., it does not provide much theoretical details. This book is also available in an online format [here](#).

- Wickham, H., & Grolemund, G. (2016). R for data science: import, tidy, transform, visualize, and model data. " O'Reilly Media, Inc."

Most of the theory I present during the course comes from these two books:

- Walpole R.E., Myers R. H, Myers S. L., and Ye K.: Probability & Statistics for Engineers & Scientists. Prentice Hall, 9th ed., 2011. Available [online](#)
- Runger G. & D. Montgomery: Applied Statistics and Probability for Engineers. Wiley, 7th ed., 2018. An old edition is available [online](#)

Additional books:

- Johnson, N.L. & Leone, F.C.: Statistics and Experimental Design Vol. 1.2, Wiley, 2nd ed., 1997.
- Draper N. & H. Smith: Applied Regression Analysis, 3rd ed. Wiley, 1998.
- Gibbons J.D.: Nonparametric Statistical Inference, Springer, 2011.

הערות

מטרת/תיאור הקורס: קורס במבוא לסטטיסטיקה וניתוח נתונים. הקורס מכסה מושגי יסוד בסטטיסטיקה (כגון מובהקות, מבחני השערות, אמידה סטטיסטית), שיטות דגימה, סוגי משתנים, מידול (כגון רגרסיה, ניתוח שונות), מבחנים א-פרמטריים. הקורס מתבסס על קורס בסיסי בהסתברות.

במסגרת הקורס נשתמש ב-R ונלמד את מימוש הפקודות הרלוונטיות באמצעות R. כמו כן, נלמד ונתרגל את כלי הויזואליזציה של R (חבילת ggplot2), וכלים שימושיים להכנת נתונים לקראת ניתוח (כדוגמת חבילות dplyr, tidyr).

במהלך הקורס יעשה שימוש ב-Data sets "פתוחים לציבור" (לדוגמה מאתר התחרויות Kaggle, מ-tidyuesday).

הקורס נועד להכין את הסטודנטים להסתכלות מוכוונת נתונים, ראייה ביקורתית, והבנה סטטיסטית בסיסית שהיא תנאי מינימלי לכל אינטראקציה עתידית שעשויה להיות להם עם נתונים שבה יתקלו בתעשייה.

ה-repository של הקורס:

https://github.com/adisarid/intro_statistics_R



--