

שם הקורס: הדמיית נתונים

שם הפקולטה: מדעי הטבע

המחלקה: מדעי המחשב

שם המרצה: ד"ר רועי יוזביץ

מתכונת הקורס: הרצאה

שנת לימודים: ג סמסטר: ב היקף שעות: 3 ש"ס נקודות זכות:

א. **מטרות הקורס** (מטרות על / מטרות ספציפיות): הקורס מציג את הספריית החשובות ביותר לעבודה עם נתונים – DATA SCIENTIST – ספריות PANDAS ו-NUMPY. כמו כן, הקורס עוסק באספקט חשוב ביותר של עבודה זו – ויזואליזציה של נתונים, הן באמצעות MATPLOTLIB והן באמצעות SEABORN. בנוסף, ילמדו הסטודנטים איך לעבוד עם נתונים המגיעים מתחרויות כדוגמת KAGGLE.

מהלך השיעורים: הקורס הינו קורס פרונטלי – הרצאה.

ב. **תוכן הקורס:**

תכנית הוראה מפורטת לפי נושאים

יחידת שיעור	נושא השיעור	הערות
1	פתיחה, מטלות ומבנה הקורס ומבוא להדמיית נתונים	
2-4	נושאים מתקדמים בספריית PANDAS	Merge, Reshape, Pivot, concantae, Mapping, Replacing, Binning and Outliers, GroupBy, Aggregation, Cross Tabulation
5-6	ויזואליזציה מתקדמת באמצעות MATPLOTLIB	
7	ויזואליזציה באמצעות SEABORN	
8	עבודה עם KAGGLE, ניקוי DATA	
9	חישוב ערכים עצמיים ווקטורים עצמיים	

	מודל PCA	10
	מודל SNE, T-SNE	11

ג. מרכיבי הציון הסופי :

עבודת גמר - 100%

ד. [ביבליוגרפיה](#): (חובה/רשות)- מסודרת לפי נושאי

הקורס.

1. Grus, Joel. *Data science from scratch: first principles with python*. " O'Reilly Media, Inc.", 2015.
2. McKinney, Wes. *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. " O'Reilly Media, Inc.", 2012.