

בית הספר למוסמכים במינהל עסקים ע"ש ליאון רקנאטי

1242.3275 - כריית נתונים למנהל עסקים עם יישומים ב-r  
Data Mining for Business Analytics with Application in r

סמסטר א' תשפ"א

**הערה:** לסטודנטים ללא ניסיון/לימוד קודם של תכנות (ברמה בסיסית לפחות) מומלץ ללמוד לפני את הקורס "טיפול יישומי בנתוני אנליטיקה עסקית".

קבוצה	יום בשבוע	שעה	תאריך בחינה	מרצה	דואר אלקטרוני	טלפון
01		מתוקשב	<a href="#">בהתאם ללוח הבחינות באתר</a>	ד"ר ענבל יהב	inbalyahav@tauex.tau.ac.il	

שעת קבלה – בתיאום מראש

### היקף הלימודים

היקף 1 י"ס לקורס  
ECTS = 4 = 2 ש"ס – European Credit Transfer and Accumulation System (ECTS), ערך הניקוד של הקורס במוסדות להשכלה גבוהה בעולם שהינם חלק מ"תהליך בולוניה".

### תיאור הקורס

מטרת הקורס הינה היכרות עם תוכנת R, ויישום מודלים סטטיסטיים מורכבים בתוכנה זו. בקורס יוצגו מספר case studies אותם ננתח בתוכנה בעזרת הידע שנצבור בשיעור. הקורס מיועד לסטודנטים המעוניינים להתמקצע ב-data analytics, ובפרט ניתוח מגמות בשוק, ניתוח לקוחות, וניתוח Big Data. הקורס ינתן בצורה מתוקשבת.

### תפוקות למידה

- עם סיום הקורס בהצלחה הסטודנט:
1. יקבל רקע נרחב בעבודה עם תוכנת R לניתוחים סטטיסטיים
  2. יקבל בסיס מקצועי לניתוח נתונים בעזרת מודלים מורכבים
  3. יקבל בסיס להמשיך להתפתח בתחום באופן עצמאי

### הערכת הסטודנט בקורס והרכב הציון

אחוז	מטלה	תאריך	גודל קבוצה/ הערות
20%	3 תרגילי בית קצרים		זוגות

## מדיניות שמירה על טווח ציונים

החל משנה"ל תשס"ט מונהגת בפקולטה מדיניות שמירה על טווח ציונים בקורסי התואר השני. עקרונות השיטה חלים על כל קורסי התואר השני, ומדיניות השמירה על טווח הציונים תיושם לגבי הציון הסופי בקורס זה. מידע נוסף בנושא זה מתפרסם בהרחבה באתר הפקולטה.

## הערכת הקורס ע"י הסטודנטים

בסיומו של הקורס הסטודנטים ישתתפו בסקר הוראה על מנת להסיק מסקנות לטובת צרכי הסטודנטים והאוניברסיטה.

## אתר הקורס

הקורס הסימסטר יינתן באופן מתוקשב דרך אתר הקורס.

הקורס מחולק לחמישה חלקים עיקריים, ושתי הרחבות, מתוכן תתבקשו לבחור הרחבה אחת בלבד. לכל חלק יוקצה בין שבוע לשבועיים להשלמתו.

בכל אחד מחלקי הקורס באתר תמצאו:

1. פורום שאלות ותשובות
2. מצגות וידאו לצפיה
3. מצגות Power point
4. מסדי נתונים לשיעור
5. קודי R לשיעור
6. תרגילים לתרגול עצמי, הכוללים פתרון. תרגילים אלו אינם להגשה.
7. תרגילי בית להגשה בזוגות/ יחידים

## תכנית הקורס \*

נושא	נושאים	פרקים בספר הלימוד וקריאת רשות
1	מבוא לתוכנות R ו-R Studio עיבוד מסד נתונים	פרקים: 1, 2 מאמרים: 5
2	ויזואליזציה של נתונים ב-R (בסיסית, ובעזרת ספריית ggplot2)	פרקים: 3
3	רגרסיה לינארית (OLS), רגרסית Log-log, סדרות זמן עיתיות	פרקים: 5, 6 מאמרים: 3 פרקים: 16, 17, 18
4	מודלי סיווג: עצי סיווג, רשתות נוירונים	פרקים: 9, 11
5	Big Data	מאמרים: 4
הרחבה	רשתות דמיון ומערכות המלצה	פרקים: 14, 19 מאמרים: 6

Shmueli G., Bruce P., Yahav I., Petal N., and Lichtendahl C. Data Mining for Business Intelligence: Concepts, Techniques, and Applications in R (2017)

1. Blei, D. M. (2012). "Probabilistic topic models." *Communications of the ACM*, 55(4), 77-84.
2. Das, Subrata. "Topics and relationship discovery from unstructured equivocal sources for situation assessment." *Information Fusion (FUSION)*, 2016 19th International Conference on. IEEE, 2016.
3. Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 963-974.
4. Ostrouchov, G., Chen, W. C., Schmidt, D., & Patel, P. (2012). Programming with big data in R. URL <http://r-pbd.org>.
5. Shmueli, G. (2010). To explain or to predict?. *Statistical science*, 25(3), 289-310.
6. Shvarts, M., Lobur, M., & Stekh, Y. (2017, April). Some trends in modern recommender systems. In *Perspective Technologies and Methods in MEMS Design (MEMSTECH)*, 2017 XIIIth International Conference on (pp. 167-169). IEEE.