

האוניברסיטה הפתוחה

20277

## **מערכות בסיסי נתונים**

חוברת הקורס קיץ 2023

כתבה: שרית סואל

יולי 2023 – סמסטר קיץ – תשפ"ג

**פנימי – לא להפצה.**

© כל הזכויות שמורות לאוניברסיטה הפתוחה.

## תוכן העניינים

א	אל הסטודנטים
ג	1. לוח זמנים ופעילויות
ד	2. תיאור המטלות
ה	3. התנאים לקבלת נקודות זכות בקורס
ו	ממ"ה 55
1	ממ"ן 11
3	ממ"ן 12
7	ממ"ן 13
11	ממ"ן 14
15	ממ"ן 15



## אל הסטודנטים,

עם הצטרפותכם ללומדים בקורס, אני מאחלת לכם הצלחה רבה, ומקווה שתמצאו בו עניין ותועלת. החוברת שלפניכם כוללת את לוח הזמנים של הקורס, תנאים לקבלת נקודות זכות ומטלות.

הקורס בסמסטר קיץ נמשך 9 שבועות בלבד ולכן יידרש ממך מאמץ ניכר לעמוד בעומס ובלוח הזמנים. חשוב להקפיד על לימוד החומר והגשת המטלות בקצב שקבענו, כדי להבטיח סיום מוצלח של הקורס. **בגלל משך הסמסטר הקצר, אין אפשרות לדחות את הגשת המטלות.**

קראו בעיון את כל הסעיפים לפני שתתחילו בלימודיכם. פרטים נוספים על המערכת המסייעת ללימוד עצמי, מרכיביה ופרטים מנהליים על ביצוע הפעילויות השונות במסגרת הלימודים, מופיעים בידיעון האוניברסיטה הפתוחה ובחוברת "מידע והנחיות הרשמה".

לקורס קיים אתר באינטרנט בו תמצאו חומרי למידה נוספים, אותם אני מפרסמת. בנוסף, האתר מהווה עבורכם ערוץ תקשורת עם צוות ההוראה ועם סטודנטים אחרים בקורס. פרטים על למידה מתוקשבת ואתר הקורס, תמצאו באתר שה"ם בכתובת:

<http://telem.openu.ac.il>

מידע על שירותי ספרייה ומקורות מידע שהאוניברסיטה מעמידה לרשותכם, תמצאו באתר הספרייה באינטרנט [www.openu.ac.il/Library](http://www.openu.ac.il/Library).

### לתשומת לב הסטודנטים הלומדים בחו"ל:

למרות הריחוק הפיסי הגדול, נשתדל לשמור אתכם על קשרים הדוקים ולעמוד לרשותכם ככל האפשר.

הפרטים החיוניים על הקורס נכללים בחוברת הקורס וכן באתר הקורס. מומלץ מאוד להשתמש באתר הקורס ובכל אמצעי העזר שבו וכמובן לפנות אלינו במידת הצורך.

אפשר לפנות אלי בכל שאלה בימי ד בין השעות 10:00-12:00, בטלפון 09-7781224. כמו-כן, ניתן גם לפנות בדואר אלקטרוני: [saritse@openu.ac.il](mailto:saritse@openu.ac.il)

בברכה,

*סרית*

מרכזת ההוראה בקורס



# 1. לוח זמנים ופעילויות (2027/ 2023)

שבוע הלימוד	תאריכי שבוע הלימוד	יחידת הלימוד המומלצת	תאריך אחרון למשלוח הממ"ן (למנחה)
1	14.7.2023-9.7.2023	פרק 1 – מבוא ; פרק 2 – מודל היחסים	
2	21.7.2023-16.7.2023	פרק 3 – מבוא לשפת SQL ; פרק 4 – SQL – רמת ביניים ; פרק 5 – SQL מתקדם	ממ"ן 11 21.7.2023
3	28.7.2023-23.7.2023 (ה צום ט' באב)	פרק 4 – SQL – רמת ביניים ; פרק 5 – SQL מתקדם	
4	4.8.2023-30.7.2023	פרק 6 – מודל ישויות-קשרים	ממ"ן 12 4.8.2023
5	11.8.2023-6.8.2023	פרק 7 – תיכון במודל היחסים	
6	18.8.2023-13.8.2023	פרק 7 – תיכון במודל היחסים ; פרק 8 – טיפוסים נתונים מורכבים ;	ממ"ן 13 18.8.2023
7	25.8.2023-20.8.2023	פרק 8 – טיפוסים נתונים מורכבים	
8	1.9.2023-27.8.2023	פרקים 9-11 – אינדקסים ; עיבוד שאילתות	ממ"ן 14 29.8.2023
9	8.9.2023-3.9.2023	פרקים 9-11 – עיבוד שאילתות חזרה	ממ"ן 15 8.9.2023

מועדי בחינות הגמר יפורסמו בנפרד

\* התאריכים המדויקים של המפגשים הקבוצתיים מופיעים ב"לוח מפגשים ומנחים".

## 2. תיאור המטלות

יש לקרוא היטב עמודים אלו לפני פתרון המטלות

פתרון המטלות הוא חלק בלתי נפרד מלימוד הקורס – הבנה מעמיקה של חומר הלימוד דורשת תרגול רב. המטלות ייבדקו על-ידי המנחה ויוחזרו לך בצירוף הערות המתייחסות לתשובות. עליך לפתור את המטלות בעצמך.

### 2.1 מבנה המטלות

כל מטלה מורכבת מכמה שאלות. בראש כל שאלה מצוין משקלה היחסי בקביעת ציון המטלה.

### 2.2 ניקוד המטלות

בקורס חמש מטלות, ומטלה מקדמת השתתפות. המטלות מלוות את חומר הלימוד של הקורס. להלן פירוט המטלות, היחידות שאליהן מתייחסת כל מטלה, וניקוד כל מטלה. משקל כל מטלה מצוין אף בראש המטלה.

משקל	חומר הלימוד	מטלה
4		מטלה (ממ"ה) 55
2	פרק 2	ממ"ן 11
4	פרקים 3, 4, 5	ממ"ן 12
6	פרקים 6, 7	ממ"ן 13
2	פרק 8	ממ"ן 14
4	פרקים 9, 10, 11	ממ"ן 15

המשקל הכולל של כל המטלות הוא 22 נקודות (משקלים). חובה לצבור נקודות במטלות במשקל כולל של 12 נקודות לפחות.

ללא צבירת 12 נקודות בהגשת המטלות לא ניתן יהיה לגשת לבחינת הגמר.

את הפתרונות למטלה עליך לרשום בצורה ברורה ומסודרת. רצוי להשאיר שוליים רחבים להערות המנחה. אם שאלה בממ"ן אינה ברורה לך, אל תהסס. להתקשר אל אחד המנחים (בשעות הייעוץ הטלפוני) לשם קבלת הסבר.



### לתשומת ליבכם:

מדיניות קורס זה היא לאשר הזנת ציון אפס במטלות שלא הוגשו כנדרש בקורס. סטודנטים אשר לא הגישו את מכסת המטלות המינימאלית לעמידה בדרישות הקורס ולקבלת זכאות להיבחן, ומבקשים שמטלות חסרות יוזנו בציון אפס, יפנו למוקד הפניות והמידע בטלפון 09-7782222 או **יעדכנו בעצמם** באתר שאילתא <http://www.openu.ac.il/sheilta>

**קורסים ⇨ ציוני מטלות ובחינות ⇨ הזנת ציון 0 למטלות רשות שלא הוגשו.**  
יש לקחת בחשבון כי מטלות אשר יוזן להן ציון אפס ישוקללו בחישוב הציון הסופי ובכך יורידו ציון זה ולא ניתן יהיה להמירן במטלות חלופיות במועד מאוחר יותר. על כן קיימת אפשרות שסטודנט אשר יעבור את הבחינה בהצלחה ייכשל בקורס (כשהממוצע המשוקלל של המטלות והבחינה יהיה נמוך מ- 60).

**כלל זה איננו חל על מטלות חובה או על מטלות שנקבע עבורן ציון מינימום.**

### 3. התנאים לקבלת נקודות זכות בקורס

- א. צבירת 12 נקודות לפחות במטלות.
- ב. ציון של 60 לפחות בבחינת הגמר.
- ג. ציון סופי בקורס של 60 נקודות לפחות.

### לתשומת לבכם!

כדי לעודדכם להגיש לבדיקה מספר רב של מטלות הנהגנו את ההקלה שלהלן:  
אם הגשתם מטלות מעל למשקל המינימלי הנדרש בקורס, **המטלות** בציון הנמוך ביותר, שציוניהן נמוכים מציון הבחינה (**עד שתי מטלות**), לא יילקחו בחשבון בעת שקלול הציון הסופי.  
זאת בתנאי שמטלות אלה **אינן חלק מדרישות החובה בקורס** ושהמשקל הצבור של המטלות האחרות שהוגשו, מגיע למינימום הנדרש.  
**זכרו!** ציון סופי מחושב רק לסטודנטים שעברו את בחינת הגמר בציון 60 ומעלה והגישו מטלות כנדרש באותו קורס

## **מטלת השתתפות במפגשים**

### **מטלה (ממ"ה) 55**

מטלת רשות

**הקורס:** 2027 – מערכות בסיסי נתונים

**סמסטר:** 2023 **משקל המטלה:** 4 נקודות

#### **למה כדאי להשתתף במפגשים?**

מטלת ההשתתפות היא מטלת רשות, והיא נועדה לעודד השתתפות במפגשי ההנחיה. זאת מתוך ההבנה שנוכחות במפגש ולמידה פעילה תורמות לפיתוח של מיומנויות אקדמיות בכלל ומיומנויות נדרשות בתחום הדעת של הקורס שלנו בפרט. מפגשי ההנחיה כוללים מידע והסברים החשובים להבנה ולהבהרה של תכני הקורס, בכלל זאת מתקיימים במפגשים דיונים ותרגולים, ולמשתתפים בהם יש אפשרות לקבל מענה לשאלות ולסוגיות שלא הובנו דיין בלמידה העצמית.

#### **בכמה מפגשים צריך להשתתף?**

המטלה מחייבת השתתפות פעילה שלך ב-70% מהמפגשים בקבוצה אליה נרשמת, לכל הפחות. כלומר:

בקבוצה בהנחייה מוגברת נדרשת השתתפות פעילה ב-7 מתוך 10 המפגשים

בקבוצה בהנחייה רגילה נדרשת השתתפות פעילה ב-5 מתוך 7 המפגשים

#### **מה נחשב להשתתפות פעילה?**

ההשתתפות תבוא לידי ביטוי בנוכחות במשך כל השיעור החי (במפגש מקוון או פרונטלי), וכן בפתיחת מצלמה (במפגש מקוון) ובתרומה לדיון בכיתה (בכתב ובע"פ) בהתאם להנחיות המנחה ובהתבסס על פעילויות שיתקיימו בשיעורים. רישום ההשתתפות הפעילה יעשה על ידי המנחה, במהלך המפגש: לעיתים בתחילת המפגש, לעיתים באמצע המפגש ולעיתים בסופו.

#### **כיצד מחושב ציון המטלה?**

השתתפות פעילה במספר המפגשים הנדרש תזכה בציון 100 על המטלה.

במצב של אי עמידה במספר המפגשים הנדרש – המטלה לא תחשב בשקלול הציון הסופי.

\*במקרים של מעבר מקבוצת לימוד אחת לאחרת – יש לעדכן את מרכז ההוראה, על מנת שניתן יהיה לאחד את המידע משתי הקבוצות.

# מטלת מנחה (ממ"ן) 11

הקורס: מערכות בסיסי נתונים

חומר הלימוד למטלה: פרק 2 – מודל היחסים

משקל המטלה: 2

מספר השאלות: 2

מועד אחרון להגשה: 21.7.2023

סמסטר: 2023

## קיימות שתי חלופות להגשת מטלות:

- שליחת מטלות באמצעות מערכת המטלות המקוונת באתר הבית של הקורס
  - שליחת מטלות באמצעות הדואר או הגשה ישירה למנחה במפגשי ההנחיה
- הסבר מפורט ב"נוהל הגשת מטלות מנחה"

לאחר אינספור סבבי בחירות הוחלט להקים בסיס נתונים המסכם את בחירות הבוחרים למועדי הבחירות השונים:

election (edate, kno)

party (pname, symbol)

running (edate, pname, chid, totalvotes)

city (cid, cname, region)

votes (cid, pname, edate, nofvotes)

בחירות (election) התקיימו בתאריך בתאריך (edate) לכנסת (kno).

למפלגה (party) יש שם (pname) והאותיות המופיעות על פתק ההצבעה של המפלגה (symbol). מפלגות רצות בבחירות השונות (running), ונרצה לדעת מי יו"ר המפלגה בבחירות אלו (chid) ומה מספר הקולות שקיבלה המפלגה בבחירות לכנסת זו (totalvotes). ערך זה מאותחל ב-0. כל מפלגה במערכת רצה לכנסת בבחירות.

נשמר מידע על היישובים השונים שבהם התקיימו הצבעות (city). ליישוב יש מזהה (cid), שם (cname) ואזור (region).

נרצה לדעת כמה הצבעות (nofvotes) קיבלה כל מפלגה בבחירות השונות לפי עיר, ומידע זה מופיע ביחס votes. אם מפלגה מסוימת לא קיבלה קולות בעיר כלשהי, היא לא תופיע ביחס עם עיר זו ולכן הערך של מספר ההצבעות (קולות) ביחס זה לא יכול להיות ריק. המערכת שומרת רק קולות כשרים. בכל בחירות היו מצביעים שהצביעו בהם.

המפתח הראשי של כל אחד מהיחסים מסומן בקו תחתון.

## שאלה 1 (50%)

נסחו את השאילתות הבאות באלגברה של יחסים. אין להשתמש בפעולת ההשמה :

- מצאו מפלגה שבעיר באר שבע קיבלה בבחירות שהתקיימו לכנסת ה-25 יותר מ-5000 קולות, והציגו גם את האותיות המופיעות על פתק ההצבעה שלה.
- מצאו זוגות של מפלגות שונות שהתמודדו באותן בחירות, ויש להן יו"ר עם אותו שם. על כל זוג מפלגות להופיע פעם אחת בלבד.
- מצאו מפלגות שהתמודדו לכנסת ה-7 ולא קיבלו אף קול (הצבעה) באזור הצפון בבחירות לכנסת זו.
- מצאו מזהה ושם של עיר שהצביעו בה בבחירות שהתקיימו ב-9.4.2019 לכל המפלגות שהתמודדו לכנסת בבחירות אלו.
- לכל עיר, מצאו מי המפלגה שקיבלה בעיר זו את מספר ההצבעות הגבוה ביותר לבחירות לכנסת שהתקיימו ב-1.11.2022. הציגו את מזהה העיר, ושם המפלגה שעונה על התנאי.

## שאלה 2 (50%)

נתונים שלושה יחסים  $r(A,B)$ ,  $s(B,C)$ ,  $t(A,C)$  :

r	
A	B
8	2
3	2
5	6
7	5
9	3
5	5

s	
B	C
6	2
2	4
8	1
8	3
2	5
6	4

t	
A	C
7	1
1	2
9	3
5	2
3	3
1	3

נתונות השאילתה הבאות באלגברה של יחסים. הציגו את יחסי התוצאה לכל אחת מהשאילתות :

- $\sigma_{B < 5}(r \bowtie s)$
- $\rho_{x(B)}(\Pi_A(r) \cap \pi_C(s)) \bowtie r$
- $\pi_{A,B}(r \bowtie s) \bowtie \pi_{A,C}(s \bowtie t)$
- $\rho_{x(A,C)}(r \cup s) \div \pi_C(t)$
- $\Pi r.B, s.B(\sigma_{r.A=s.B \wedge r.B < s.C}(r \times s))$

# מטלת מנחה (ממ"ן) 12

**הקורס:** מערכות בסיסי נתונים

**חומר הלימוד למטלה:** פרקים 3, 4, 5 – שפת SQL

**משקל המטלה:** 4

**מספר השאלות:** 4

**מועד אחרון להגשה:** 4.8.2023

**סמסטר:** ג2023

**שימו לב:**

- את המטלה יש להגיש באמצעות מערכת המטלות המקוונת באתר הבית של הקורס בלבד
- הסבר מפורט ב"נוהל הגשת מטלות מנחה"

בממ"ן זה עליכם להגדיר במערכת לניהול בסיסי נתונים PostgreSQL, בסיס נתונים עבור בחירות הבוחרים במועדי הבחירות השונים כפי שתואר בממ"ן 11. כמו כן יהיה עליכם לאכלס אותו בנתונים ולבצע בו פעולות שונות. להלן תזכורת לתבניות היחסים:

election (edate, kno)

party (pname, symbol)

running (edate, pname, chid, totalvotes)

city (cid, cname, region)

votes (cid, pname, edate, nofvotes)

**המפתח הראשי של כל אחד מהיחסים מסומן בקו תחתון.**

באתר הקורס בלשונית SQL מופיע קובץ הנחיות כלליות לממ"ן 12. הנחיות אלו תקפות למטלה וכן לבחינה. בנוסף, בסוף המטלה ישנן הנחיות כלליות הנוגעות למטלה. עיינו בקובץ זה ובהנחיות הכלליות למטלה בטרם תתחילו בפתרון המטלה.

**המשימות:**

א. (10 נק') באתר הקורס בלשונית SQL מופיע קובץ עם הגדרת טיפוסים נתונים מתאימים לתכונות היחסים שהוגדרו בתחילת המטלה. הגדירו יחסים אלו בבסיס הנתונים באמצעות שאילתות תוך שימוש בטיפוסים הנתונים בדף העזר לכל שדה, הגדירו מפתחות ראשיים, מפתחות זרים וכן אילוצים נוספים שיש להגדיר עבור יחסים אלה מתוך תיאור המערכת כפי שמופיע בממ"ן 11.

- ב. (10 נק') כתבו הדק (טריגר) שבהוספה או עדכון של הצבעות למפלגה מעדכן את סך הקולות שהמפלגה קיבלה בבחירות אלו. יש לכתוב את הטרigger בצורה יעילה, כך שלא תתבצע קריאה של יחס שלם.  
שם הטרigger יהיה T1 ושם פונקציה הטרigger `trigfl()`.
- ג. (10 נק') טענו נתונים לכל אחד מהיחסים באמצעות שאילתות הוספה. צריכות להיות לפחות 3 רשומות בכל אחד מהיחסים. נתונים לדוגמה יש בקובץ העזר למטלה באתר.
- ד. (70 נק', 10 נק' לסעיף)  
נסחו את השאילתות הבאות ב-SQL ושמרו אותן.  
בסעיפים אלו **אין להשתמש** בפונקציות או במבטים שאתם יוצרים.  
אם יש צורך, ניתן להשתמש בפונקציה `cast (A as B)` הממירה את תכונה A לטיפוס B.  
ניתן להשתמש ב-`with` אחד לכל היותר בכל סעיפי השאלה, ורק אם באמת יש בו צורך, ז"א שהיחס שהוגדר ישמש לפחות פעמיים בשאילתה עצמה.  
אין לקנן שאילתות ב-`from`.  
אם יש יותר מרשומה אחת שעונה על תנאי השאילתה, יש להציג את כולן.
- מצאו לכל מפלגה שהתמודדה בבחירות בתאריך 2.3.2020 כמה קולות קיבלה בעיר ryde end (יש להציג רק מפלגות שקיבלו קולות בעיר זו).
  - חשבו לכל מפלגה שהתמודדה בבחירות לכנסת ה-3 את כמות הקולות שקיבלה בכל מחוז בנפרד בבחירות לכנסת זו.
  - מצאו ערים שמעולם לא הצביעו בהן למפלגת 'life party'. הציגו את שם העיר ואת האזור שלה.
  - מצאו בחירות שהתמודדו בהן מספר המפלגות הגבוה ביותר. הציגו את תאריך הבחירות ולאילו כנסת הן היו.
  - מבין המפלגות שרצו לכנסת השלישית ומעולם לא קיבלו קולות בערים הממוקמות בהרים (hills) מצאו את המפלגה שקיבלה את סכום הקולות הנמוך ביותר בכל הבחירות שהיא התמודדה בהן.
  - מצאו את המפלגה השנייה בגודלה בכנסת ה-3, והציגו את שמה.
  - מצאו זוגות של מפלגות שרצו בכל מערכת בחירות ביחד. ז"א שהמפלגה הראשונה רצה בכל מערכות הבחירות שהמפלגה השנייה רצה, וגם להיפך. על כל זוג מפלגות להופיע רק פעם אחת בתוצאה.

### הנחיות כלליות:

- יש להגיש את המטלה בקובץ יחיד הכולל את כל הסעיפים יחד **לפי הסדר**. יש להגיש קובץ `sql` יחיד בתוך קובץ `zip` או `rar`.
- לפני כל סעיף יש להוסיף הערה עם מספר הסעיף, למשל: `--a`: לסעיף א, ו-`--d4` לסעיף ד. אין להוסיף מלל נוסף. ה: `--` מסמן הערה.

- מערכת PostgreSQL לא שומרת את השאילתות הרגילות, אלא רק פונקציות וטריגרים. בתחילת פתרון המטלה שמרו את תשובתכם באמצעות save as המאפשרת שמירת השאילתה כקובץ .sql. לקובץ זה הוסיפו את הסעיפים הבאים. מומלץ לשמור בסוף כל סעיף.
- בפונקציות ובטריגרים, יש להגדיר שהם נכתבים בשפת PL/pgSQL, ועל כן, בסוף הפונקציה או הטריגר יש להוסיף LANGUAGE plpgsql.
- בהגדרת פונקציה, לאחר ה-as ולפני הגדרת השפה יש להוסיף \$\$ כדי לציין שזהו גוף הפונקציה.
- שימו לב, בפתרון המטלה יש להשתמש רק בחומר הלימוד של הקורס. **אין** להשתמש בפונקציות נוספות כגון פונקציות windows למיניהן שעשויות לעבוד עם המערכת.
- יש להגיש את המטלה **במערכת המטלות**. יש לצרף את קובץ ה-.sql. ששמרתם.





# מטלת מנחה (ממ"ן) 13

הקורס: מערכות בסיסי נתונים

חומר הלימוד למטלה: פרקים 6, 7 – מודל ישויות קשרים ותיכון בסיס הנתונים

משקל המטלה: 6

מספר השאלות: 5

מועד אחרון להגשה: 18.8.2023

סמסטר: 2023

## קיימות שתי חלופות להגשת מטלות:

- א. שליחת מטלות באמצעות מערכת המטלות המקוונת באתר הבית של הקורס
  - ב. שליחת מטלות באמצעות הדואר או הגשה ישירה למנחה במפגשי ההנחה
- הסבר מפורט ב"נוהל הגשת מטלות מנחה"

## שאלה 1 (18 נקודות)

- א. (12 נק') שרטטו דיאגרמת ישויות קשרים לתיאור בסיס הנתונים המסכם את בחירות הבחורים במועדי הבחירות השונים כפי שהיה נתון בממ"ן 11, כולל כל האילוצים הנובעים מהנתונים.
- ב. (6 נק') כיצד תשתנה הדיאגרמה אם מפלגה יכולה לרוץ רק במערכת בחירות אחת? עדכנו את הדיאגרמה (רק את החלק הרלוונטי לשינוי) ואת היחסים אם יש עדכון ביחסים.

## שאלה 2 (25 נקודות)

- נרצה לתכנן מערכת מידע למוזיאון. המוזיאון מציג תערוכות לפי נושאים בגלריות השונות. לגלריה נשמר שמה ומיקומה במוזיאון. נרצה לדעת לתערוכה בין היתר מהו נושא התערוכה, ומי אוצר התערוכה (curator). כל תערוכה ממוקמת בגלריה אחת. בכל תערוכה מוצגים פריטים ולפריט נרצה לדעת בין היתר מהו שמו, מי האמן היוצר, את שנת היצירה ותיאור קצר על הפריט. חלק מהפריטים הם תמונות, ונרצה לדעת מה הם ממדי התמונה (אורך ורוחב), והפריטים האחרים הם פסלים, ונרצה לדעת מה גובהם ומשקלם.
- א. (10 נקודות) שרטטו דיאגרמת ישויות-קשרים לתיאור מערכת זו. הוסיפו תכונות לטיפוסי הישויות ולטיפוסי הקשרים לפי הצורך. אל תשכחו לסמן מידות ריבוי, אילוצי השתתפות ומפתחות ראשיים.
  - ב. (10 נקודות) לכל אחת מהטענות הבאות קבעו אם היא נכונה או לא לפי הדיאגרמה בלבד, ונמקו את טענתכם.
    1. ניתן להציג פריט בכמה תערוכות.
    2. יכולה להיות תערוכה שאין בה פריטים.
    3. בגלריה מוצגת לכל היותר תערוכה אחת.
    4. יכול להיות פריט שאינו תמונה ואינו פסל.

- ב. יכולה להיות גלריה שלא מוצגות בה תערוכות.  
 ג. (5 נקודות) המירו את הדיאגרמה ששרטטתם בסעיף א לתבניות יחסים וסמנו את המפתח הראשי בכל תבנית יחס.

### שאלה 3 (10 נקודות)

נתונה מערכת היחסים הבאה הכוללת 4 תבניות של יחסים:

$R(\underline{A}, B, C)$

$S(\underline{A}, \underline{D}, E)$

$T(A1, A2, \underline{D}, C, G)$

$U(\underline{C}, F)$

המפתח הראשי של כל אחד מהיחסים מסומן בקו תחתון.  
 התכונות  $A1, A2$  הן מאותו טיפוס נתונים כמו התכונה  $A$ .  
 שרטטו דיאגרמת ישויות קשרים לתיאור המערכת.

### שאלה 4 (23 נקודות)

נתונה תבנית יחסים  $R=(A, B, C, D, E, G)$ , ואוסף תלויות פונקציונליות  $F$  החלות עליה:  
 $F=\{B \rightarrow E, C \rightarrow E, GE \rightarrow CD, BD \rightarrow C, D \rightarrow G, BE \rightarrow AB\}$

- א. מצאו כיסוי קנוני ל- $F$ .  
 ב. מצאו את כל המפתחות הקבילים של  $R$ .  
 ג. (8 נקודות) נתון פירוק של  $R$  לשני יחסים:

$R1=(C,D,G,E)$

$R2=(A,B,C,E)$

- ג.1. (2 נקודות) האם הפירוק משמר מידע? נמקו.  
 ג.2. (2 נקודות) האם הפירוק משמר תלויות? נמקו.  
 ג.3. (4 נקודות) קבעו באיזו צורה נורמלית נמצאים כל אחד מהיחסים שבפירוק  $3NF, BCNF$  או אף אחד מהם). נמקו בקצרה את תשובתכם.  
 ד. פרקו את  $R$  פירוק משמר מידע לפי אלגוריתם הפירוק ל- $BCNF$  ליחסים שכולם ב- $BCNF$ .  
 האם הפירוק שהתקבל משמר תלויות? נמקו.

### שאלה 5 (24 נקודות)

נתונה תבנית יחסים  $R=(A, B, C, D, E)$  ואוסף תלויות פונקציונליות  $F$ , החלות עליה:  
 $F=\{B \rightarrow D, C \rightarrow B, C \rightarrow BE, BCD \rightarrow E, ABD \rightarrow CE\}$

- א. חשבו כיסוי קנוני ל- $F$ .

- ב. מצאו את כל המפתחות הקבילים של  $F$ .
- ג. פרקו את  $R$  פירוק משמר מידע ומשמר תלויות בצורה הגבוהה ביותר האפשרית. לכל אחד מהיחסים ציינו באיזו צורה נורמלית הוא נמצא.
- ד. הוסיפו תלות אחת לאוסף התלויות המקורי כך שהיחס יהיה ב-BCNF. נמקו מדוע היחס ב-BCNF.



# מטלת מנחה (ממ"ן) 14

הקורס: מערכות בסיסי נתונים

חומר הלימוד למטלה: פרק 8 – טיפוס נתונים מורכבים

משקל המטלה: 2

מספר השאלות: 4

מועד אחרון להגשה: 29.8.2023

סמסטר: 2023

## קיימות שתי חלופות להגשת מטלות:

- שליחת מטלות באמצעות מערכת המטלות המקוונת באתר הבית של הקורס
  - שליחת מטלות באמצעות הדואר או הגשה ישירה למנחה במפגשי ההנחיה
- הסבר מפורט ב"נוהל הגשת מטלות מנחה"

## שאלה 1 (40 נקודות)

את סעיפים א, ב, ג-יש להגיש מוקלדים בקובץ עיבוד תמלילים כלשהו (כגון קובץ txt או קובץ doc), ואין להגישם כתמונה.

בשאלה 2 בממ"ן 13 התבקשתם לשרטט דיאגרמת ישויות קשרים לתיאור מערכת מידע למוזיאון. שאלה זו מתייחסת למערכת הדומה מערכת שבשאלה אך אינה זהה לה.

בגלריית galileo art הממוקמת בקומה 3 בבניין 5 של המוזיאון מתוכננת כרגע תערוכה אחת: תערוכת "כוכבים". התערוכה עוסקת בכוכבי לכת, האוצרת שלה היא "דניאלה ארץ" ויש בה 2 פריטים: ציור בשם "נוגה" של האמנית "נוגה ס" שאורכו 200 ס"מ ורוחבו 100 ס"מ, וציור עגול בשם "מאדים" של האמן "מ אדום" שקוטרו 125 ס"מ.

א. ב-XML, תוך שימוש בייצוג מקוון מתאים.

ב. ב-JSON, עם אותו ייצוג מקוון

ג. באמצעות שלשות RDF.

ד. כגרף RDF.

## שאלה 2 (20 נקודות)

נתון RDF לייצוג חלקי של המערכת הנתונה בממ"ן 11 :

k18 instanceof kneset.	p30 pname "elm".
k17 instanceof kneset.	p30 symbol "el".
k18 edate "27.3.2023".	r2 instanceof running.
k17 edate "20.2.2019".	r2 chid "456".
r1 instanceof running.	r2 totalvotes "4545454".
r1 chid "123".	k18 kin r2.
r1 totalvotes "676767".	p30 pin r1.
k18 kin r1.	k17 kin r3.
p22 pin r1.	r3 instanceof running.
p22 pname "cherry".	r3 chid "789".
p22 symbol "ch".	r3 totalvotes "98989".
p22 instanceof party.	p22 pin r3.
p30 instanceof party.	

כתבו את 2 השאילות הבאות ב-SPARQL ביחס ל-RDF הנתון :

א. מצאו שם מפלגה ושם יו"ר של המפלגה עבור מפלגות שרצו בבחירות שהתקיימו ב-

27.3.2023.

ב. הציגו לכל כנסת את שמות המפלגות שהתמודדו בבחירות לכנסת זו. התוצאה תכיל את

מזהה הכנסת ואת שם המפלגה.

## שאלה 3 (20 נקודות)

נתון מאגר של 10 מסמכים הקשורים למערכת בחירות. נרצה לחשב TF-IDF עבור 2

מהמסמכים : d1 ו-d2 עבור שאילתה שרוצה למצוא את היו"ר של מפלגת "השולחנות". נתונה

השאילתה :

Query: tables, chairman

ידועים הנתונים הבאים : במסמך d1 יש 250 מילים, ובמסמך d2 יש 150 מילים.

המילה tables מופיעה במסמך d1 35 פעמים, ובמסמך d2 10 פעמים.

המילה chairman מופיעה במסמך d1 15 פעמים, ובמסמך d2 30 פעמים.

המילה tables מופיעה ב-4 מסמכים והמילה chairman ב-6 מסמכים.

חשבו באמצעות TF-IDF את הרלוונטיות של כל מסמך לשאילתה, וקבעו איזה מסמך ידורג גבוה

יותר. הראו את חישובי הביניים. הציגו 3 ספרות לאחר הנקודה העשרונית בחישוב.

#### שאלה 4 (20 נקודות)

נתונה מערכת הכוללת 3 דפים לדירוג.

דף 1 מצביע לדף 2.

דף 2 מצביע לדף 1 ולדף 3.

דף 3 מצביע לדף 1.

כמו כן, נתון כי  $\delta=0.18$ .

חשבו את דירוג שלושת הדפים האלו לפי אלגוריתם PageRank לאחר כל אחת מ-2 האיטרציות

הראשונות, וקבעו מי יהיה הדף בדירוג הנמוך ביותר. הציגו 3 ספרות לאחר הנקודה העשרונית

בחישובים.





# מטלת מנחה (ממ"ן) 15

הקורס: מערכות בסיסי נתונים

חומר הלימוד למטלה: פרקים 9-11 – אינדקסים ועיבוד שאילתות

משקל המטלה: 4

מספר השאלות: 5

מועד אחרון להגשה: 8.9.2023

סמסטר: 2023

## קיימות שתי חלופות להגשת מטלות:

- שליחת מטלות באמצעות מערכת המטלות המקוונת באתר הבית של הקורס
  - שליחת מטלות באמצעות הדואר או הגשה ישירה למנחה במפגשי ההנחיה
- הסבר מפורט ב"נוהל הגשת מטלות מנחה"

## שאלה 1 (22 נקודות)

א. (10 נק') בנו עץ  $B+$  עבור סדרת הערכים הבאה (משמאל לימין) כאשר  $n=4$ . בפיצול צומת, אם ניתן לפצל את הערכים שווה בשווה, יש לפצל אותם שווה בשווה. אחרת, מספר הערכים הקטן יהיה בצומת השמאלי.

22, 127, 24, 109, 14, 58, 78, 60, 5, 36, 48, 412, 97, 66, 86, 100, 120, 30.

ב. (8 נק') בטלו את הערכים הבאים מהעץ שקיבלתם בסעיף א אחד אחרי השני (משמאל לימין). הראו את העץ המתקבל לאחר כל מחיקה.

24, 89, 30, 100

ג. (4 נק') כמה ערכים יש להוסיף לעץ שקיבלתם בסעיף א כדי שהוא יגדל ברמה? נמקו.

## שאלה 2 (22 נקודות)

א. (8 נק') בנו מבנה של גיבוב סטטי עבור 10 הערכים הראשונים (משמאל לימין) המופיעים בשאלה 1 סעיף א (מהערך 22 עד לערך 36). הניחו כי בכל סל נכנסות עד 2 רשומות אינדקס, ופונקציית הגיבוב היא  $h(x) = x \bmod 8$ .

ב. (10 נק') בנו מבנה של גיבוב מתרחב עבור אותה סדרת ערכים מסעיף א, הניחו כי בכל סל נכנסות עד 3 רשומות, וכי פונקציית הגיבוב היא  $h(x) = x \bmod 16$  כאשר ערך הפונקציה מיוצג בבינארי.

ג. (4 נק') בהינתן רוחב תחילית גלובלית ששווה ל-11, לכל היותר כמה כניסות במדריך (directory) יכולות להצביע לסל אחד? נמקו בקצרה.

### שאלה 3 (9 נקודות)

נתון יחס לדוגמה לטבלת city הנתונה בממ"ן 11. בנו מבנה של מפת סיביות עבור העמודה region.

cid	cname	region
22	ryde end	north
77	east strat	south
33	grandetu	hills
88	royalpre	hills
11	carlpa	hills
44	lommont	north
66	grand sen	south
99	kinged haven	hills
55	el munds	south

### שאלה 4 (24 נקודות)

נתונים שלושה יחסים  $r(A, B, C)$ ,  $s(C, D, E)$ ,  $q(E, F)$ . כמו כן ידועים הנתונים הסטטיסטיים הבאים:

$$n_r=20,000 \quad n_s=30,000 \quad n_q=10,000$$

$$V(A,r)=15,000 \quad V(B,r)=500 \quad V(C,r)=20,000 \quad V(C,s)=20,000$$

$$V(D,s)=2,000 \quad V(E,s)=9,000 \quad V(E,q)=8,000 \quad V(F,q)=1,000$$

תחום הערכים של כל אחת מתכונות הערכים הוא מספרים טבעיים.  
ניתן להניח שלתכונות עם שם זהה ערכים זהים, עד כמה שניתן.  
העריכו את מספר השורות בכל אחד מהביטויים הבאים.

א.  $s \bowtie \sigma_{B=85}(r)$

ב.  $s \bowtie q$

ג.  $\Pi_A(r) \cap \Pi_F(q)$  הביטוי

נמקו את כל תשובותיכם.

### שאלה 5 (23 נקודות)

ידועים הנתונים הבאים עבור בסיס הנתונים לסיכום בחירות הבחורים במועדי הבחירות השונים הנתון בממ"ן 11:

$$n_{city} = V(cid, city) = V(cid, votes) = V(cname, city) = 150$$

$$n_{votes} = 400,000$$

$$V(region, city) = 5$$

בנוסף, ידוע כי רק ב-2% מהמקרים, מפלגה קיבלה פחות מ-20 קולות ביישוב.  
נתונה השאילתה הבאה באלגברה של יחסים:

$$\Pi_{cname, pname}(\sigma_{city.cid=votes.cid}(\sigma_{region='north'}(city) \times \sigma_{nofvotes < 20}(votes)))$$

- א. (5 נק') נסחו בעברית מה מחשבת השאילתה.
- ב. (10 נק') העריכו (בממוצע) את מספר השורות בתוצאת השאילתה. נמקו את תשובתכם.
- ג. (8 נק') אם ידוע שבסיס הנתונים מכיל מידע על 20 מערכות בחירות, בכל מערכת בחירות רצות עד 10 מפלגות, ומספר הערים הוא כפי שנתון בתחילת השאלה, כמה שורות לכל היותר ולכל הפחות יכולות להיות ביחס votes? הסבירו את משמעות תשובותיכם בבסיס נתונים זה.