

## פרויקט קורס חלק ב' – PHP

הנחיות להגשת התרגיל:

1. תאריך הגשה 26/1/20 בשעה 23:55

2. יש להגיש שני קבצים: קובץ pdf וקובץ zip עם כל קבצי האתר. קובץ ה-pdf צריך להכיל עמוד שער עם שמות ותעודות הזהות של המגישים, וכתובת האינטרנט של האתר שבניתם. את קובץ ה-pdf יש להכניס לתוך קובץ ה-zip. פורמט שם קובץ ה-zip הינו: id1\_id2.zip.
3. את הקובץ יש להגיש דרך אתר הקורס במקום המתאים ב-moodle על ידי **אחד** מבני הזוג (לא לשכוח לציין בתוך פתרון תרגיל הבית את שם המשתמש של בן הזוג אשר בחשבונו שמורים קבצי האתר).
4. איחור בהגשת התרגיל יגרור קנס בגובה 20% מהציון עבור כל יום איחור (פרט למקרים חריגים כגון מילואים. במקרים אלה יש לפנות טרם הגשת התרגיל)

בחלק זה של הפרויקט עליכם להשלים את אתר המערכת אשר מודלה בחלק הראשון של הפרויקט, על פי ההוראות המפורטות בעמודים הבאים.

בכל אחד מהסעיפים ישנן הנחיות עבור עמוד מסוים באתר המערכת.

העבודה תתבצע מול מסד נתונים מרכזי. פרטי החיבור הם:

שם המסד: <your campus name>

שרת: techniondbcourse01.database.windows.net

שם משתמש: <your campus name>

סיסמא: Qwerty12!

להלן תיאור הדפים השונים:

### 1. דף הבית

דף הבית של האתר צריך לכלול את הפרטים הבאים:

- כותרת לאתר המערכת.
- תמונה שמתאימה למערכת/לחברה (על האתר להיות נגיש לכולם ולדעת להציג טקסט חלופי במקרה שהתמונה לא עולה).
- קישורים לשאר העמודים אשר מתוארים בגליון.
- תיאור כללי של מטרת המערכת.

המשך תיאור הדף בעמוד הבא


- טבלה המכילה את הנתונים הסטטיסטיים הבאים :

- a. מזהה המשתמש עם כמות הביקורות הגדולה ביותר.
- b. מזהה המשתמש שסכום מספר המשתמשים שביקורותיו עזרו להם הוא המקסימלי. במקרה של שוויון יש להציג את מזהה המשתמש הקטן ביותר לפי יחס סדר לקסיקוגרפי.
- c. מזהה המשתמש אשר ממוצע היחס בין מספר המשתמשים שהביקורת שלו עזרה להם ומספר המשתמשים הכללי אשר קראו את הביקורת (ממוצע מעל כל הביקורות) הוא הגבוה ביותר. במקרה של שוויון יש להציג את מזהה המשתמש הקטן ביותר לפי יחס סדר לקסיקוגרפי. עבור נתון זה יש להתעלם מביקורות אשר לא נקראו בכלל.
- d. מזהה המוצר בעל מספר הקשרים הגבוה ביותר עם מוצרים אחרים (שימו לב למבנה הרלציה המתוארת בהמשך הקובץ – מדובר ביחס חד צדדי: מוצר 'א' (reference\_asin) ביחס עם מוצר 'y' (related\_asin)). במקרה של שוויון יש להציג את מזהה המוצר הקטן ביותר לפי יחס סדר לקסיקוגרפי.

דוגמא לתבנית של דף בית :

## Bamazon Site

A short description on the system....



[Load Data](#)  
[Add Review](#)  
[Display Visual Data](#)

General Data			
Highest amount of reviews	highest sum of helpful ratings	Highest average ratio	Highest number of relations

## 2. טעינת נתוני המערכת מקובץ

הקישור הראשון יוביל לעמוד אשר מאפשר טעינה של שני קבצים שונים (קבצי CSV).  
הקבצים אשר יש לתמוך בטעינתם הם:

- 1) קובץ ביקורות למוצרים.
  - 2) קובץ יחסים בין מוצרים.
  - 3) קובץ המכיל מילים אשר ישמשו ליצירת ביקורת אוטומטית.
- בתרגול מספר 7 (בנושא PHP) ניתן לראות כיצד טוענים קובץ למסד נתונים (עליכם להתאים את טעינת הנתונים למסד כפי שלמדנו בתרגול).
- הקישור צריך להכיל כפתור שאמצעותו המשתמש יוכל לבחור את הקובץ שברצונו להעלות למסד נתונים.
- לאחר טעינת הנתונים יש להקפיץ הודעה למשתמש המבשרת האם טעינת הנתונים עברה בהצלחה (יש להציג את מספר השורות אשר נכנסו למסד בהצלחה מתוך מספר השורות הכולל).

מצורפים לתרגיל הקבצים הבאים:

- 1) reviews.csv אשר מכיל מידע אודות הביקורות למוצרים השונים
- 2) relations.csv אשר מכיל מידע אודות קשרים בין מוצרים במערכת.
- 3) corpus.csv אשר מכיל מילים שישמשו ליצירת ביקורת אוטומטית.

דוגמא לעמוד:

**Load Data**

Press on the relevant button in order to load data of a specific file

**Reviews Data**

Choose your file:

Choose File No file chosen submit

**Relations Data**

Choose your file:

Choose File No file chosen submit

**Corpus Data**

Choose your file:

Choose File No file chosen submit

### 3. טופס הזנת ביקורת אוטומטית

הקישור השני יוביל לטופס אשר יאפשר הוספת ביקורת אוטומטית חדשה למערכת.

השדות אשר הטופס צריך להכיל הם :

- 1) מזהה משתמש – מחרוזת (יש לחייב את המשתמש למלא ערך זה).
  - 2) מזהה המוצר – מחרוזת אשר מציינת את מזהה המוצר (יש לחייב את המשתמש למלא ערך זה).
  - 3) מספר המילים אשר הביקורת תכיל – שדה מספרי (מספר שלם בתחום [1,100]). במקרה שלא הוכנס ערך, נתייחס למספר זה כ-100.
  - 4) הסתברות לבחירת מילה – שדה מספרי (מספר ממשי בתחום (0,1]). במקרה שהוכנס הערך 0 או שלא הוכנס ערך כלל, יש להופכו לערך '1'.
  - 5) ציון למוצר – שדה מספרי (מספרים שלמים בתחום [1,5], כאשר 1 הוא הציון הנמוך ביותר שניתן לתת למוצר). ערך ברירת מחדל לשדה זה הינו 1.
  - 6) כפתור שליחת טופס.
  - 7) כפתור מחיקת תשובות קיימות.
- הטופס יפעל באופן הבא :

לאחר שכלל הנתונים הוזנו – על המערכת לעבור על המילים אשר נמצאות ברלציה corpus (הסבר על הרלציה מופיע בסוף התרגיל). עבור כל מילה נגדיל מספר ממשי רנדומלי בתחום [0,1]. המספר הנ"ל מציין את הסף לשימוש במילה בביקורת, כלומר אם הערך שהוגרל גדול מההסתברות להכנסת מילה לביקורת אשר הוזנה בטופס – נבחר שלא להכניס את המילה לביקורת (אחרת – נכניס). יש להכניס מילים לביקורת עד אשר מגיעים למספר המילים הרצוי (אשר הוזן בטופס). מילים שונות יופרדו על ידי תו רווח יחיד. כמו כן אם המעבר על המילים ברלציה הסתיים אך מספר המילים בביקורת קטן ממספר המילים אשר הוזן, יש לבצע מעברים נוספים – עד אשר מגיעים לספר המילים הרצוי. לאחר יצירת הביקורת, יש להוסיף את הרשומה לטבלת reviews.

על מנת להגריל מספר ממשי – ניתן להשתמש בפונקציה המתוארת בעמוד הבא :

<https://www.php.net/manual/en/function.rand.php>

**שימו לב** – על מנת להכניס ביקורת למסד עליכם להכניס את נקודת ציון הזמן שבה הביקורת נקלטה במערכת. ניתן לקרוא אודות המשתנה בקישור הבא (בחלק אשר מתאים ל-sql server):

<https://database.guide/how-to-return-the-unix-timestamp-in-sql-server-t-sql/>

**המשך תיאור בעמוד הבא**

**בונוס –** הסבירות שפלט הביקורת לפי השיטה הנתונה תהיה בעלת משמעות כלשהי הינה נמוכה מאוד. על כן, **בונוס** לפתרון של הסעיף הנ"ל, ניתן לנסות לייצר באופן אחר ביקורות באיכות גבוהה יותר. למגיש הבונוס יינתנו עד ל-10 נק' בונוס לפי החלטת סגל הקורס. יש להגיש קובץ PDF אשר מתאר את השיטה הנבחרת, כמו כן, יש לכלול את פלט השיטה ברלציה שנקראת makesSenseReview (המבנה שלה יהיה זהה לרלציה reviews) – יש לייצר 5 ביקורות שונות. הנקודות יינתנו בהתאם להערכת השיטה והפלט על ידי סגל הקורס.

דוגמא לעמוד :

## Add New Review

Fill this form in order to add new review

### Information

User ID	<input type="text"/>	*
Product ID	<input type="text"/>	*
Number of tokens	<input type="text"/>	
Probability	<input type="text"/>	
Rating	<input type="text"/>	

#### 4. הצגת מידע חזותי – ביקורות מוצר

בקישור האחרון יש להציג מידע חזותי אודות ביקורות על מוצר נבחר. עליכם לייצר טופס אשר תומך בהכנסת מזהה מוצר והמידע הרצוי להצגה, אשר נבחר מבין האפשרויות הבאות:

(1) מספר אנשים אשר קראו את הביקורת.

(2) מספר האנשים אשר הביקורת עזרה להם.

(3) הציון אשר ניתן לביקורת.

לאחר שליחת הטופס (הוכנס מזהה מוצר ונבחר מידע להצגה) יוצג גרף אשר מתאר את הנתונים הדרושים עבור המוצר. על הנתונים להיות ממוינים לפי חתימת זמן בסדר עולה ועבור כל מוצר אין להציג מעל ל-10 נקודות בגרף (יש להציג את 10 הנקודות הראשונות).

לדוגמא, הוכנס מזהה מוצר 'x' ונתון להצגה "מספר אנשים אשר קראו ביקורת" – עבור הקלט הנ"ל יוצג גרף אשר מציג את 10 הביקורות הראשונות (יתכן ויהיו פחות ביקורות) ועבור כל ביקורת מציג את מספר האנשים אשר קראו אותה. על הגרף להציג את הנתונים בסדר ממוין לפי חתימת הזמן. על כותרת הגרף להכיל את מזהה המוצר הנבחר ובנוסף יש להציג במקרא הגרף את הנתון אשר נבחר (כמו בדוגמא המצורפת).

על מנת לייצר גרפים באמצעות PHP אנו מצרפים לקבצי התרגיל את התיקיה pChart. עליכם לצרף את התיקיה לקוד הפרויקט על מנת שהוא יעבוד (יש לבצע push לקבצי התיקיה על מנת להעלותם לgit repository). התיקיה הנ"ל מכילה מחלקות אשר תומכות ביצירת הגרפים. כמו כן, מצורף לתרגיל קובץ קוד מתועד אשר מדגים כיצד ניתן לייצר גרף. שם הקובץ הוא pChart\_example.php.

דוגמא לטופס לפני שליחה:

## Visual data page

Please enter the required details

Product ID	<input type="text" value="X"/>
<input checked="" type="radio"/>	Total number of readings
<input type="radio"/>	Number of 'helpful' votes
<input type="radio"/>	Overall rating
<input type="button" value="Submit"/>	

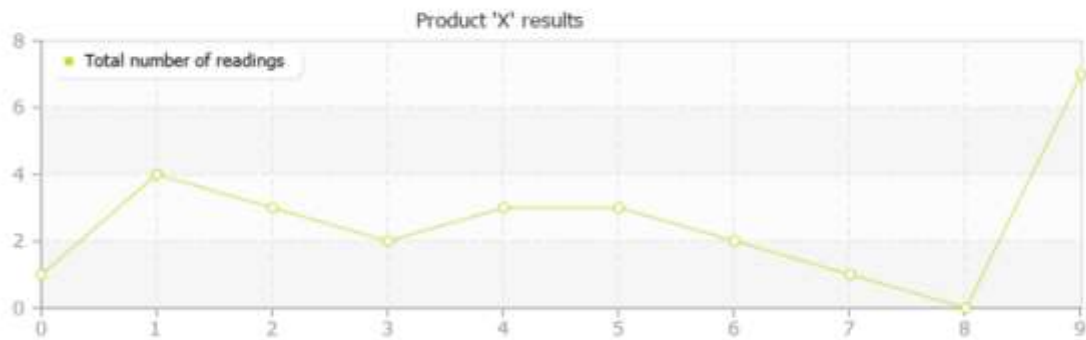
דוגמא לגרף הנבחר לאחר שליחת הטופס :

## Visual data page

Please enter the required details

Product ID

☐ Total number of readings  
☐ Number of 'helpful' votes  
☐ Overall rating



## פרטים נוספים :

התרגיל מורכב מהקבצים הבאים :

**(1 reviews.csv)** - מכיל מידע אודות הביקורות למוצרים השונים.

שדות הקובץ הם :

i. reviewrID – מזהה משתמש אשר מזין את הביקורת.

ii. Asin – מזהה המוצר אשר עליו נכתבת הביקורת.

iii. reviewText – תוכן הביקורת.

iv. unixReviewTime – שדה אשר מתאר את חתימת הזמן שבה הביקורת מקלטה במערכת.

v. overall – הציון אשר המשתמש נתן למוצר.

vi. helped – מספר האנשים אשר שצינו כי הביקורת עזרה להם.

vii. overallRead – מספר האנשים הכולל אשר קראו את הביקורת.

**(2 relations.csv)** - מכיל מידע אודות קשרים בין מוצרים במערכת.

שדות הקובץ הם :

i. reference\_asin – מזהה מוצר אשר נמצא ביחס עם מוצר אחר.

ii. related\_asin – מזהה מוצר אשר reference\_asin נמצא ביחס עמו.

iii. type – טיב היחס.

**(3 corpus.csv)** - מכיל את המילים השונות אשר משמשות ליצירת תגובות אוטומטיות.

שדות הקובץ הם :

i. token – מילים המשמשות ליצירת התגובה האוטומטית.

**(4 createTables.sql)** – מכיל סקריפט ליצירת טבלאות המסד. הטבלאות זהות מבחינה

סכמטית למבנה הקבצים המתואר.

בהצלחה!