**עבודת סיכום –BIG DATA**

**30% מדעי המחשב**

העבודה בא לסכם נושאי לימוד הקשורים בשימוש ב-Big Data

כל תלמיד חייב להגיש עבודה אישית שמתבסס על מאגר מידע גדול וינסה להציג את המידע באופן משמעותי תוך השערת השערה ולנסות לבדוק האם על סמך מדגם גדול, יהיה ניתן להוכיח את ההשערה.

התלמיד יחקור תופעה שבו אספו ממספר גדול של מקורות מידע וריכזו יחד למאגר מידע.

**תיק הפרויקט**

התיק יחולק ל-4 חלקים שמשקפים את העבודה שהתלמיד השקיע

**פרק 1:** נושא החקר. התלמיד יציג את הרקע לפרויקט עם סיבה למה בחר את הנושא. התלמיד יציג את ההנחות המרכזיות של הפרויקט בקווים מאוד כלליים.

**פרק 2:** תיאור המסד נתונים של העבודה. העבודה תכיל סקירה של העמודות של הטבלה ומידע ראשוני על מבנה המאגר מידע, האם יש יותר ממאגר אחת וגם יש לתאר אותו.

**פרק 3:** שיטות המחקר. חוזרים לפרק הראשון. אם בפרק ראשון הצגנו את השאלות המעניינים אותנו, כאן נוסיף את הקוד שיש לכתוב בפייתון כדי שנוכל לחשב ולהציג את הגרפים שיקלו על הבנת הנתונים

**פרק 4:** סיכום העבודה הכולל הצגת המידע לאחר עיבוד נתונים ראשונים והאם הנתונים אכן מצביעים על מה שהנחנו נכון או לא. לבסוף נוסיף כמה שורות רפלקציה על מהלך כתיבת הפרויקט

הניקוד:

ספר הפרויקט :

פרק ראשון : 10% מהציון

פרק שני : תיאור הנתונים 10%

פרק שלשי : שיטות מחקר: האלמנטים שהשתמשתם בפייתון הכוללים: הוספת עמודות, בניית DataFrames ו- Series חדשים לפי הדרישות. שימוש בקיבוץ לפי קרטריונים: (groupby) וחיבור שתי קבוצות :( Merge), סינון שורות לפי קריטריונים 30%

סיכום העבודה: הצגת גרפים , התאמה בין בחירה האמצעי לבין המידע. שילוב כמה שיטות להצגת נתונים 15%

הצגת העבודה 15%

מטלות בזמן הבחינה: 20

בונוס 10% על תחכום