מיקוד לבחינת הגמר – ויזואליזציה של מידע לתלמידי מנע״ס

סמסטר 2025ב, כל המועדים

|  |  |  |  |
| --- | --- | --- | --- |
| מפגש | תוכן עקרוני | הסברים והערות | צפי למספר שאלות |
| 1-2 | התקנת Tableau, חיבור לנתוני CSV, הגדרה של שדות ותרשימים בסיסיים, כלי תרשים ראשוניים – כולל דשבורד ראשוני (כולל התאמת גודלו ל-Automatic)  הגדרה של Row ו-Columns  התפקיד של Marks כולל Colors הכי נפוץ, לצרכי פילוח, וגם Labels תוויות נתונים  שימוש בפילטרים  יצירת תרשים עמודות בסיסי, המרת תרשים קו לעמודות - כולל דרך הגדרה ספציפית, וכולל תוך שימוש ב-Automatic ושינוי סוג המשתנה מבדיד לרציף ולהפך (Continuous, Discrete), והמשמעות של הגדרת סוג תרשים ב-Marks  מיון ערכים ויצירת דשבורד  ההבדל בין Measure (כגון מכירות) ל-Dimension (כגון סוג מוצר).  **רוב מפגש 2 עסק בתרגולים של הגדרות אלו, בסגנון מבחן.** | שאלות בסיסיות שכוללות תיאור מקרה עם נתונים ומשתנים מסויימים, והמטרה היא לענות / לדון בהיגדים שקשורים לאופן הגרירה / ההגדרה של משתנים והשימוש בכלים השונים כבסיס ליצירת התרשים הנדרש  למרות שהמבחן ״על נייר״ – כדאי להתאמן על הנתונים כדי להרגיש את זה | 2-3 |
| 2 | שדות מחושבים – Calculated Fields  היכולת להגדיר משתנים חדשים על בסיס פעולות חישוביות ו/או שדות אחרים לטובת הניתוח  שימוש ב-Bins: קיבוץ ערכים של משתנה כמותי ״יחד״ – למשל – קיבוץ כל מרווח של 5 שנים יחד (אפשרי רק לערכים מספריים). משתני Bin הם תמיד משתנים בדידים (Discrete). | בהחלט תתכנה שאלות שבהם תדרשו להפעיל מחשבה האם שימוש ב-Calculated Field ו/או ב-Bins הוא מתאים לפי סוג הנתונים והמטרה | 1-2 |
| 3,6,7 | שימוש בכלי AI ליצירת נתונים (הנקניק של אור): היכולת ליצור קבצי Excel באמצעות ה-Chat וחיבור מתאים ל-Tableau  יצירה של דשבורד על ידי גרירת תרשימים ואלמנטים נוספים (כגון התמונה של אור שושן, אתרי אינטרנט ועוד) וייצוא / שמירה בענן של Tableau Public  היכרות עם GitHub לשיתוף ואחסון.  הצורך ביצירת index.html בתור התחלה כולל באגים וזיהויים  היכרות עם תפריט ה-Pages לשם פרסום האתר | השיעור הזה הוא ככל הנראה הבסיסי ביותר / הטכני ביותר: רוב השאלות שעשויות להשאל עליו קשורות לשימוש בכלי המתאים / מושגי | 1 |
| 4 | מושג ה-Quick Table Calculation  כלי שמאפשר התאמת ערכים בתרשים למטרה מוגדרת. ספציפי – חישוב מובנה (סוג של פונקציה) שמופעל על נתונים / תרשים קיים כולל עיבודו האוטומטי (דירוגים, ערכים מצטברים, סטיית תקן ועוד) | צריך להכיר את סוגי החישובים בתוך Quick Table Caclulation  שהוצגו ספציפית במערך השיעור.  כלומר: Running Total, Difference, Percent of Total, Percent of Difference | 2 |
| 5 | פרשנות של מידע ויצירת דשבורדים נוספים: מצב שבו קיימים גיליונות רבים בקובץ Excel, שאנו עוקבים ומפרשים את המהות של כל אחד מהם, ואז מייצרים באמצעות Tableau מיזוג / איחוד שלהם (שימוש בכלי ה-Join). ההבדל בין Join לבין תרשימי ERD כולל נראות התפריטים | סוג אחד של שאלות – הכרות בסיסית מאד עם הכלי ומטרתו. סוג שני של שאלות – להכיר את סוגי ה-Join (Inner, Outer וכו׳ ומשמעותם) | 2 |
| 8 | ניתוח קשרים בין שני משתנים על בסיס סקדלבלוט (ScatterPlot) ורגרסיה לינארית (Linear Regression)  כולל תהליכי עבודה רלוונטיים לחישובי תשואות (מניות), בניית ה-Excel הרלוונטי, שמירת הנתונים כ-Data ולא כנוסחה, חיבור ל-Tableau של האקסל הקשור, שינוי ערכים רלוונטיים ל-Dimension כדי לקבל את תרשים הפיזור (כולל השגיאה שיכולה להיווצר במידה ולא עושים זאת). הוספת קו מגמה ומשמעותו, כולל פרשנות ערכיו, בעיקר בהקשר ל-Pvalue ומשוואת הקשר הלינארי | השאלות יכולות להיות מגוונות, גם לגבי הכלי עצמו, גם לגבי הצגה של תרשים לקוי והסיבות שיכולות להוביל לכך, לרבות פרשנות הממצאים הסטטיסטיים | 2-3 |
| 9 | רגרסיה מרובה – יצירת משוואת קשר בין מספר מסבירים למוסבר אחד יחיד. ביצענו זאת באמצעות Analysis Toolpak (אללייסיס טולפה): הצורך להגדיר את המשתנים הרלוונטיים, המרת נתונים טקסטואליים למספריים – כולל שתי אפשרויות לעשות כן (בצורה מדורגת במשתנה אחד, או במשתני דמי רבים וההקשר לסוג הנתון לצורך הכרעת הטיפול)  החשיבות של שימוש בערכים מספריים עם נתונים  הצורך בשמירה על נתונים היסטוריים בתהליך העיבוד  שימוש בכלי Analyze Data ב-Excel (שעובד רק באקסל כששפת הממשק אנגלית)  שימוש ב-Analysis Toolpak ובכלי ה-Regression גם לטובת קשר מרובה משתנים, | הצורך ברגרסיה מרובה בניתוח תופעות מורכבות, שימוש ב-Pvalue ומשמעות ממצאיו, ההיגיון בנתונים והמשמעות במקרה של תקלות (בהתאם לנלמד) | 2-3 |
| 10 | מיני מרתון – שאלות חזרה בלבד בסגנון בחינה | מבחן לדוגמא! עם תשובות! | |
| 11 | מודלים לוקאליים – LLM  הצורך ב-LLM לוקאלי, יתרונותיו, חסרונותיו, עקרון הקוד הפתוח, יצרניות עיקריות (בהתאם למה שהוצג במחברת), יתרונות כגון התאמות המודל, כוונון שלו, אבטחת מידע, שימוש בהתנסות לוקאלית ככלי להבנה משמעותית יותר של העולם הזה. לרבות נתוני טבלה משווה בעמ׳ 120 במחברת. לא צריך לשנן בעל פה הגדרות של דרישות חומרה, כן צריך להכיר באופן כללי, את היעילות הגבוהה שיש ל-GPU לטובת יישומי LLM, והצורך בזכרון RAM משמעותי.  הכלי LM Studio, בתור כלי GUI עם ממשק משתמש גרפי להתקנת LLM מקומי, אפשרויות מיון ובחירה במודלים מתאימים למערכת על בסיס חיווי, ההבדל בין מודלים קטנים וגדולים והיכולת להפעיל מודלים קטנים גם על גבי פחי אשפה.  הערכת טיב הביצוע של LLM מקומי גם על בסיס משימות והגדרות פשוטות יחסית (בדיקת פלט ראשוני). וההגדרות בעמ׳ 130: tok/sec, QAT, GGUF, MLX, Vision enabled  כלי ה-ollama להתקנה של LLM מקומי בגישת CLI | כמובן שלא תצטרכו להשיב לשאלות טכניות עמוקות, אלא להכיר את הכלי, את אופן ההתקנה, את המגבלות, את הרכיבים המשמעותיים ביותר המשפיעים על טיב התפקוד, וברמת ההגדרות הטכניות העמוקות – יש לכם הסבר מלא בעמ׳ 130. | 3-5 |
| סה״כ | | | 20 |