**README**

בקובץ זה מוגשים הקבצים הבאים כחלק מהפרויקט לניתוח מזג האוויר באזור Central-West Brazil:

1. :big\_exercise.docx מסמך מרכזי ומעמיק שמתאר את שלבי העבודה בתרגיל זה. המסמך כולל:
   * מקורות הנתונים וגודלם (מעל 11 מיליון שורות).
   * הסבר על טעינת הנתונים למסד DuckDB ושימוש ב SQLite - לטבלאות קטנות.
   * השאלות המרכזיות שעליהן עונה הניתוח.
   * דוגמאות לנתונים מעובדים שהוצגו בדאשבורד.
   * תיאור הסיפור שמספר המחקר ואופן בנייתו.
   * ויזואליזציות שנבחרו להצגת הממצאים.
   * מדריך טכני להרצת הדאשבורד ופתרון בעיות נפוצות.
2. central\_west\_hourly\_weather\_sample.csv: קובץ בפורמט .csv המקביל לקובץ הנתונים הגדול רק עם 500 שורות ראשונות.
3. :duckdb\_and\_sqlite.py קובץ זה מכיל קוד Python שמבצע את השלבים הבאים:

* טעינת נתונים ל:DuckDB - הקובץ מטעין קובץ CSV גדול למסד הנתונים DuckDB ויוצר ממנו טבלה מרכזית.
* עיבוד וניתוח נתונים: הקובץ מבצע שאילתות SQL להפקת תובנות מהנתונים, כמו ממוצע טמפרטורה לאורך זמן, סך המשקעים, ותנאים קיצוניים לפי תחנות.
* יצירת טבלאות קטנות ב:SQLite - התוצאות של השאילתות נשמרות כטבלאות קטנות ב SQLite - שימוש בדאשבורד.

1. central\_west\_data.db: מסד הנתונים בפורמט SQLite, הכולל את הטבלאות הקטנות ששימשו לניתוח הנתונים והוויזואליזציות בדאשבורד.
2. streamlit\_dashboard.py: קובץ הקוד הראשי של הדאשבורד שנכתב ב-Streamlit. הקוד מחולק לעמודים עם ניווט צדדי ומכיל את כל הוויזואליזציות, הטבלאות והתובנות המרכזיות מהנתונים.
3. requirements.txt: קובץ המכיל את כל חבילות הפייתון הנדרשות להרצת הדאשבורד, כולל Streamlit, Matplotlib, Seaborn, Pandas, ו-WordCloud.