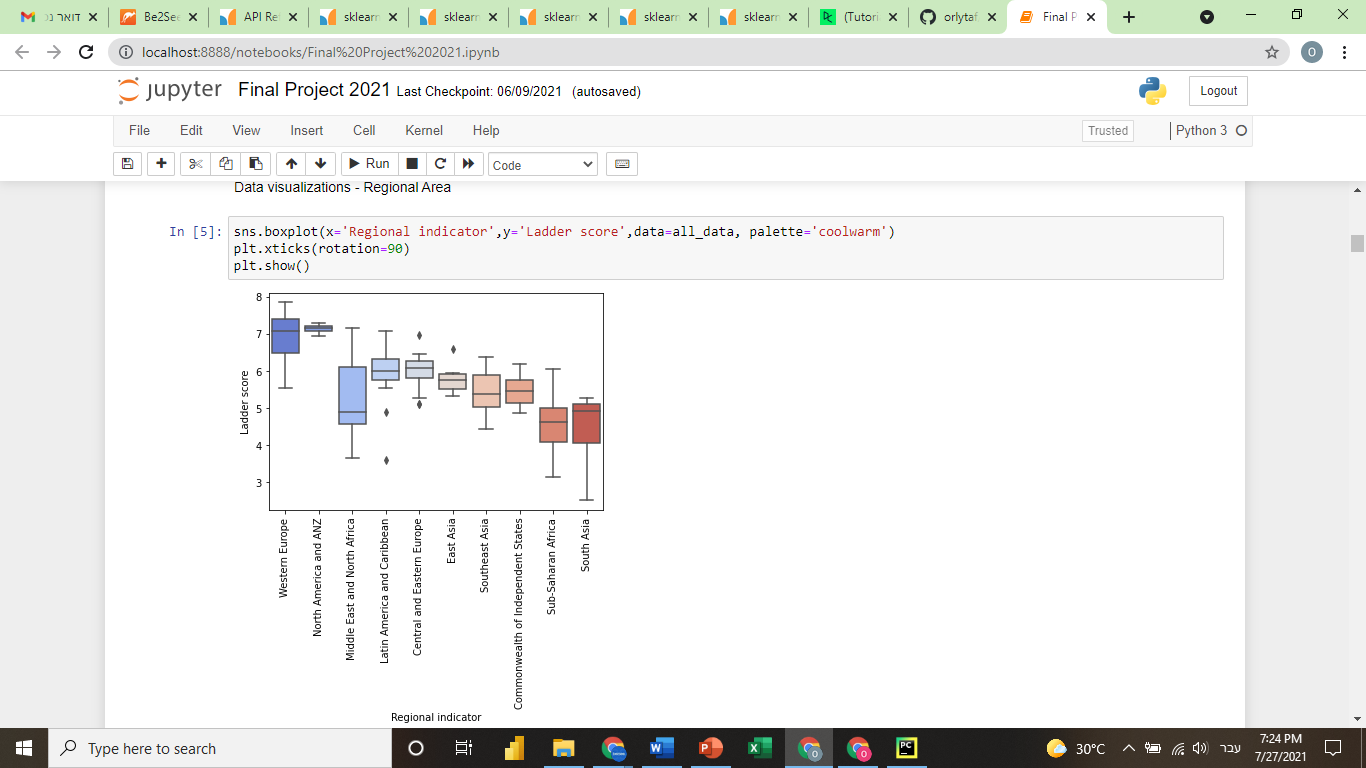
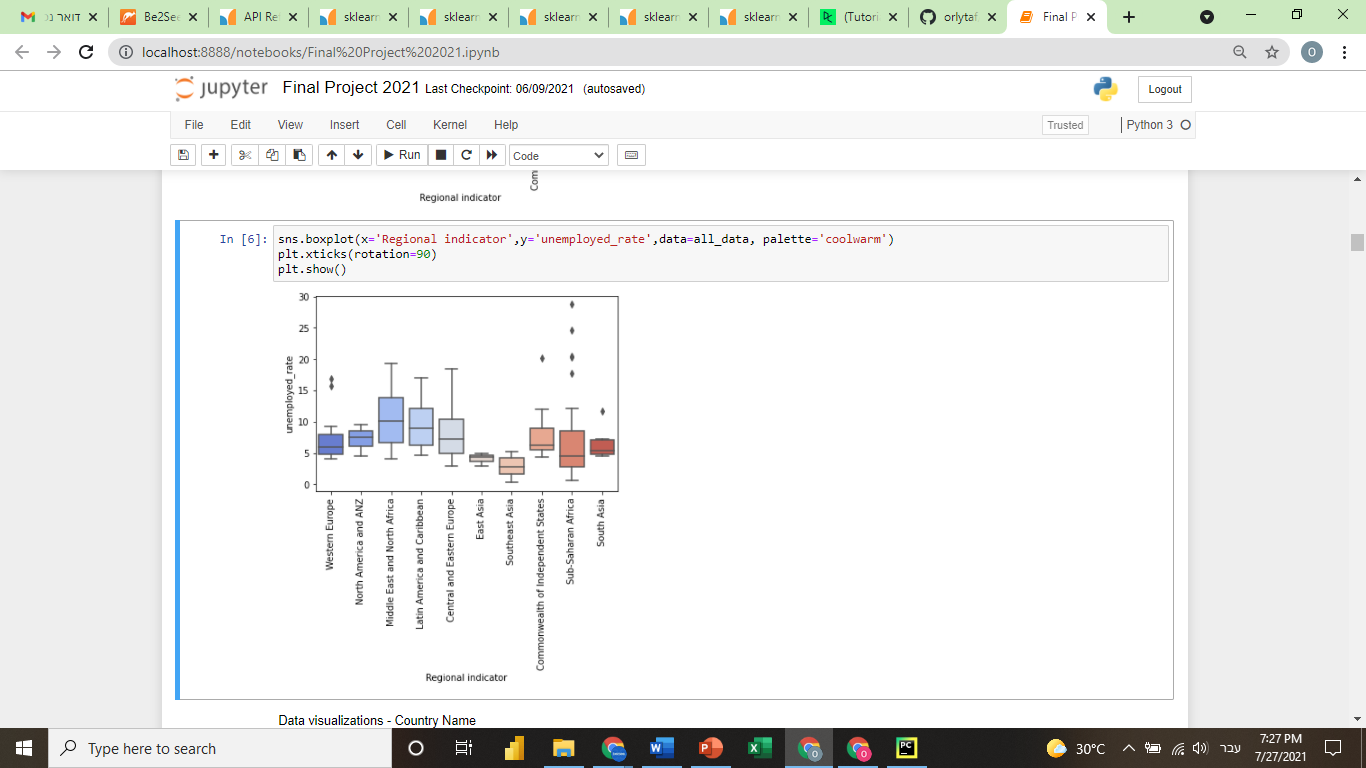
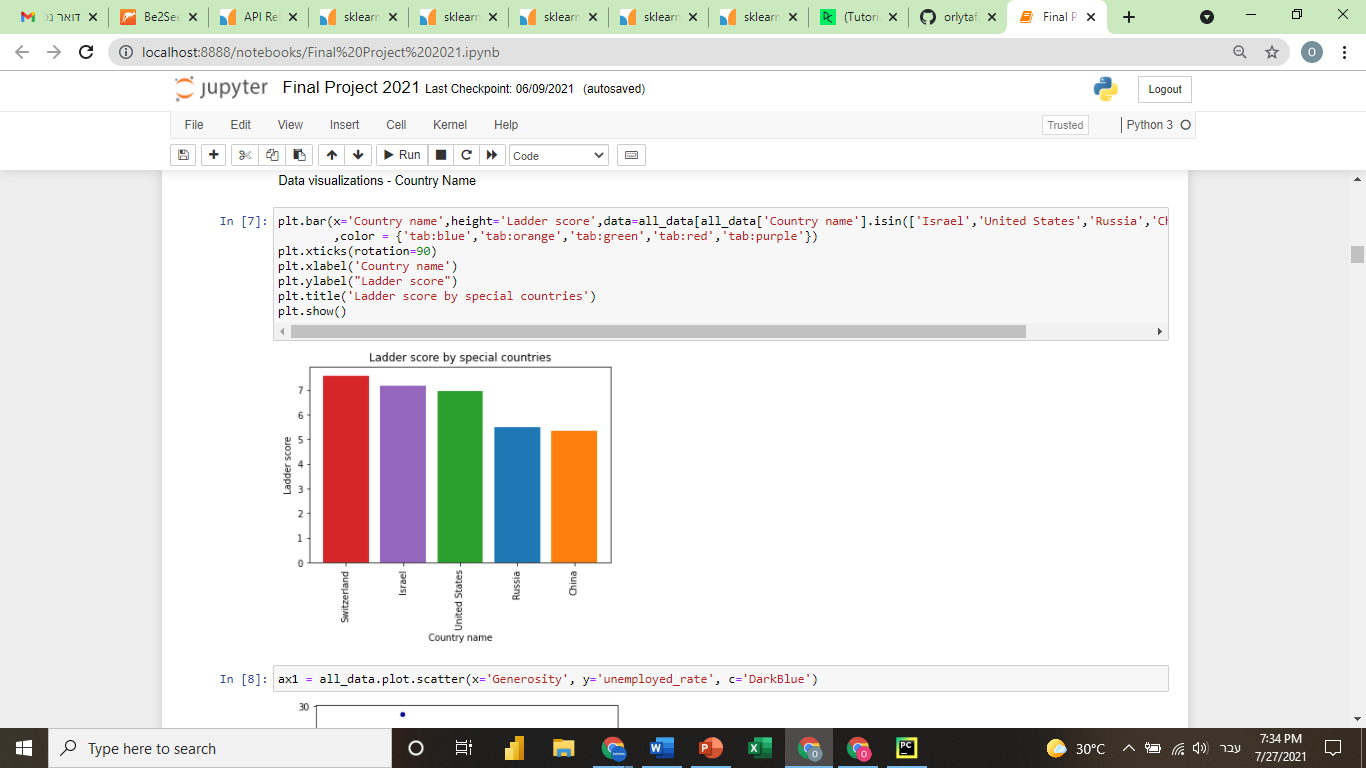
**לפני שאתחיל להציג את הפרויקט אוכל לספר שהוספתי מספר עמודות ממקורות חיצוניים ל-דאטא פריים שקיבלתי, וזו הפעם הראשונה שאני עוסקת בפייתון ובמודלים של ML ולא היה לי רקע ראשוני לפני כן.**

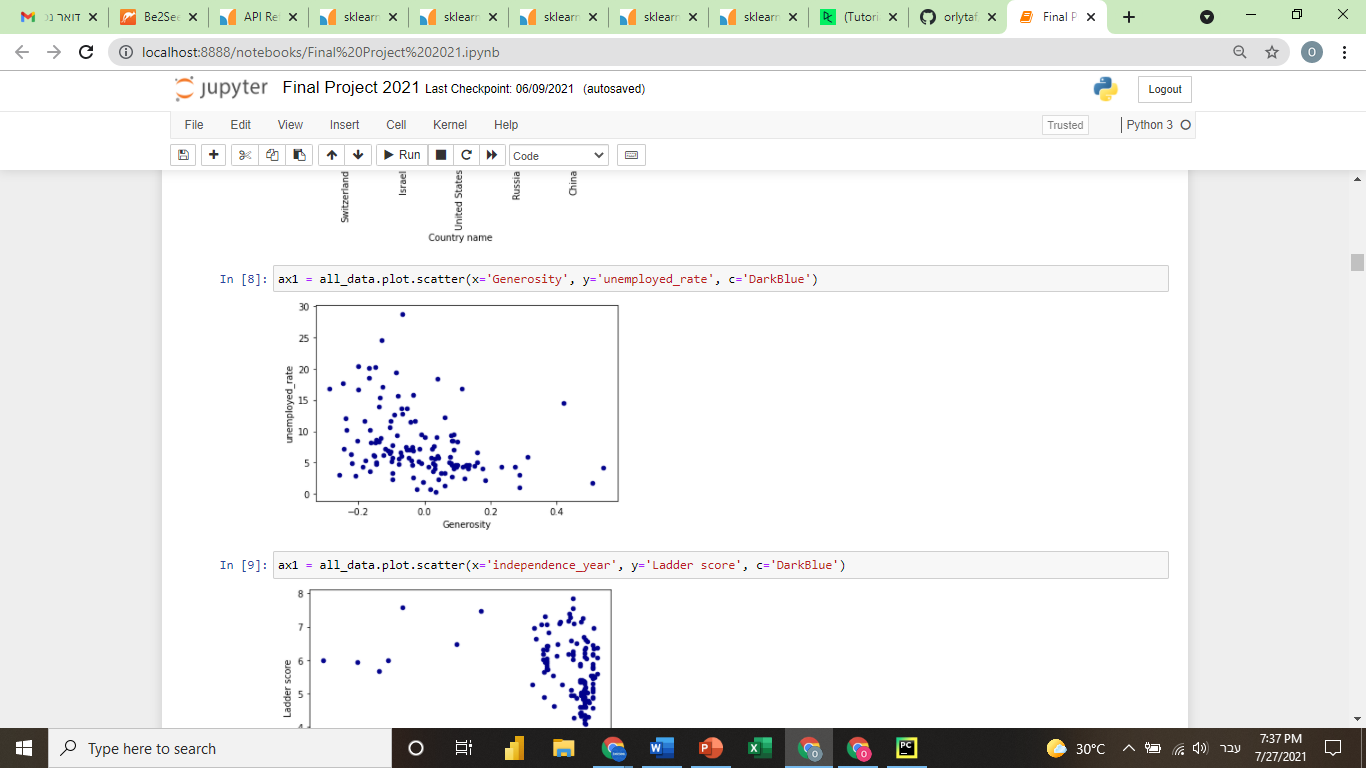
1. **ויזואליזציות לפי איזור:**

  
  
**מסקנות מהגרף:** באיזורי מערב אירופה וצפון אמריקה אוסטרליה וניו זילנד (2 ה-box-ים משמאל) יש את ציוני מדד האושר הגבוהים ביותר.

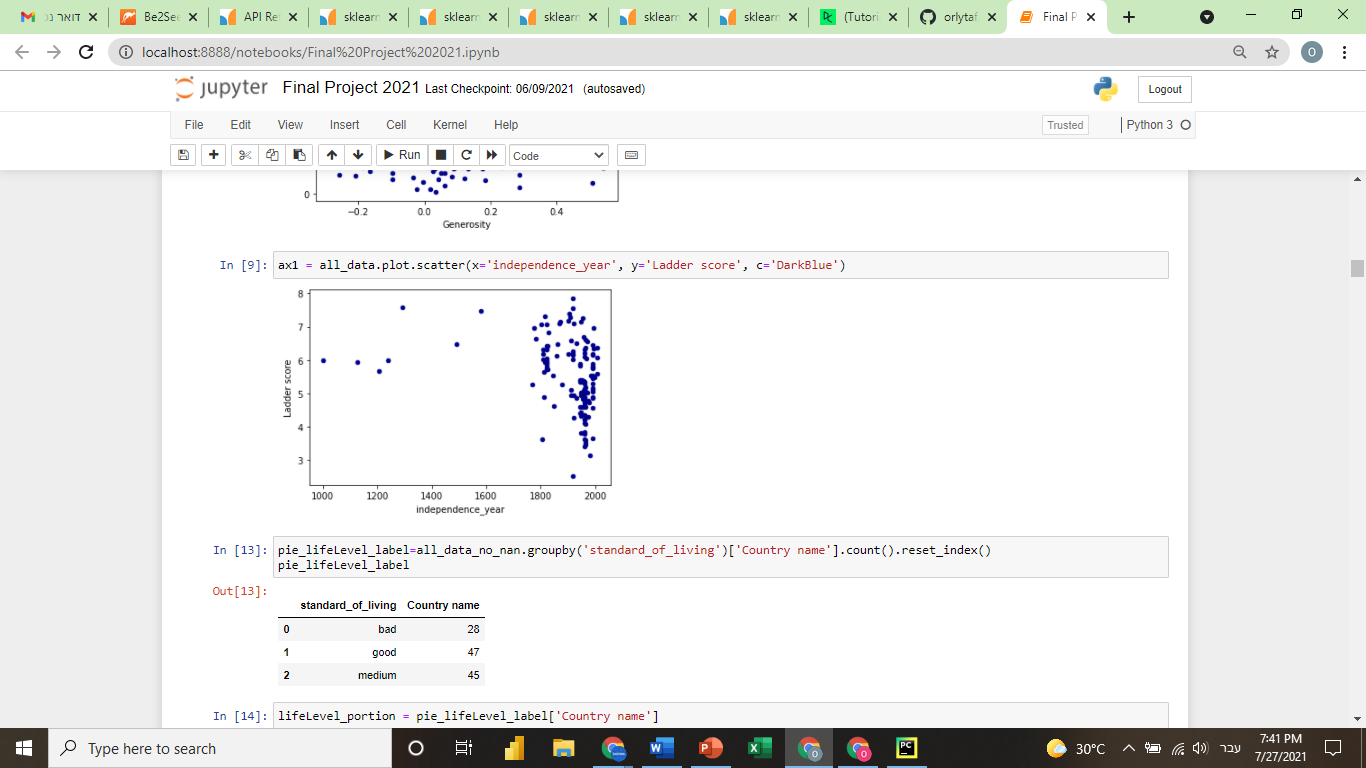


**מסקנות מהגרף:** באיזור צפון אפריקה והמזרח התיכון שיעור האבטלה הממוצע הוא הגבוה ביותר מבין כל האיזורים. באיזור דרום מזרח אסיה ממוצע שיעור האבטלה הוא הנמוך ביותר. ובאיזור אפריקה שמדרום למדבר סהרה יש הרבה חריגות בשיעורי האבטלה.   
  
**2. ויזואליזציות לפי מדינה:**

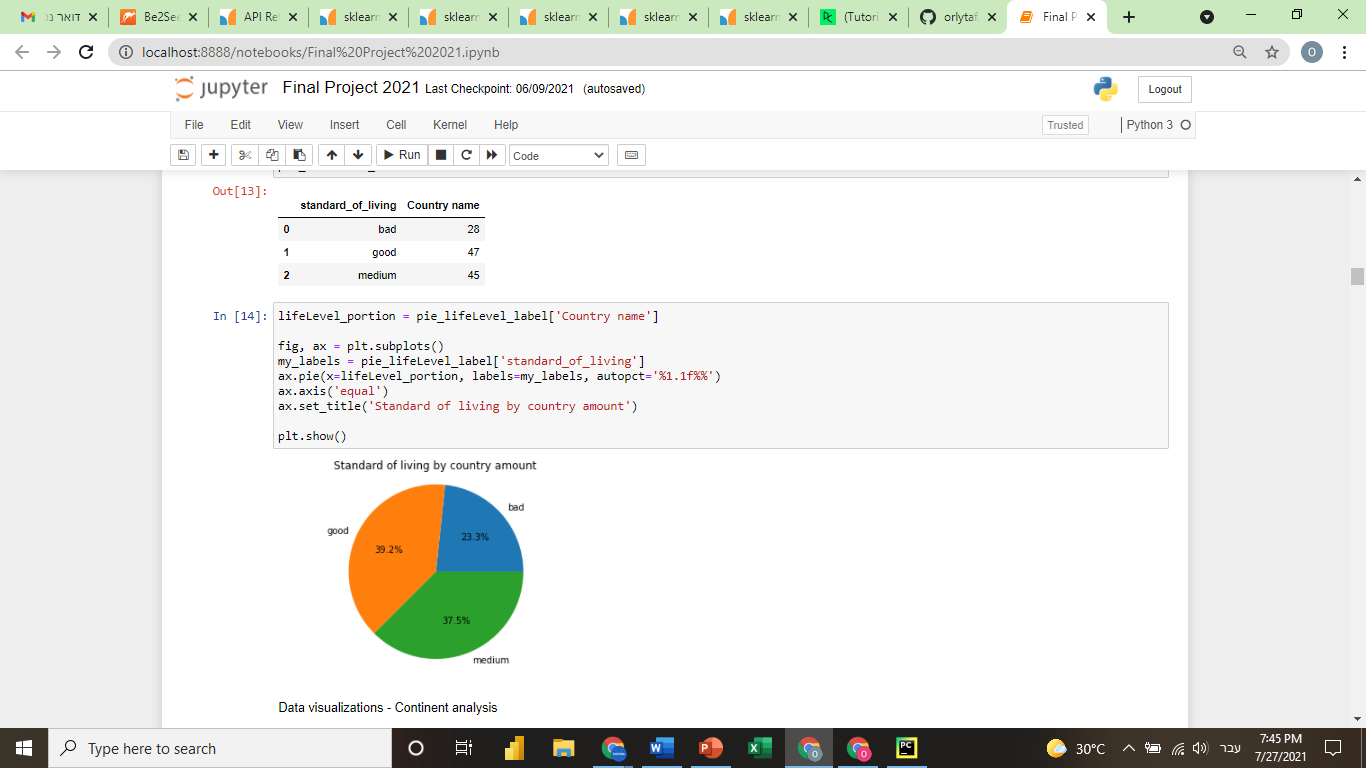
  
**מסקנות מהגרף:** זהו ***bar plot*** (תרשים עמודות) שבו לקחתי 5 מדינות שעניין אותי לחקור מהם ציוני מדדי האושר שלהן, זו ביחס לזו. מהתוצאות יצא שהמובילה במדד האושר היא שוויץ, מיד אחריה מופיעה מדינת ישראל ומי שמדד האושר שלה היה הכי נמוך היא סין.



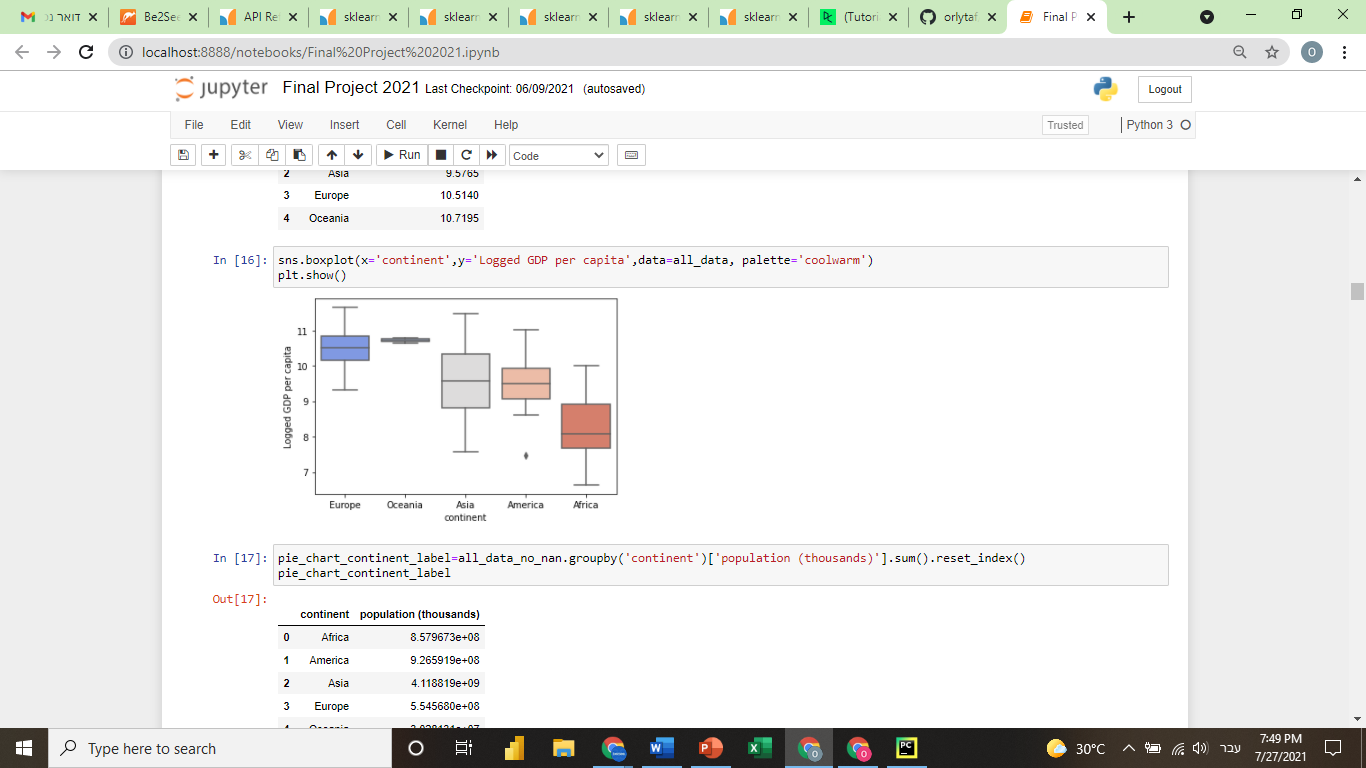
**מסקנות מהגרף:** זהו ***תרשים פיזור*** שבו נמדד שיעור האבטלה (שהוספתי בצורה עצמאית לדאטה פריים שלי) ביחס לפרמטר הנדיבות שהיה קיים במאגר הנתונים. ציפיתי שיתקיים קשר בין 2 הפרמטרים האלו וחשבתי שאמצא שככל ששיעור האבטלה יהיה נמוך יותר רמת הנדיבות תהיה גבוה יותר. אולם בפועל התרשים לא מעיד על קשר וודאי בין 2 הפרמטרים האלו.



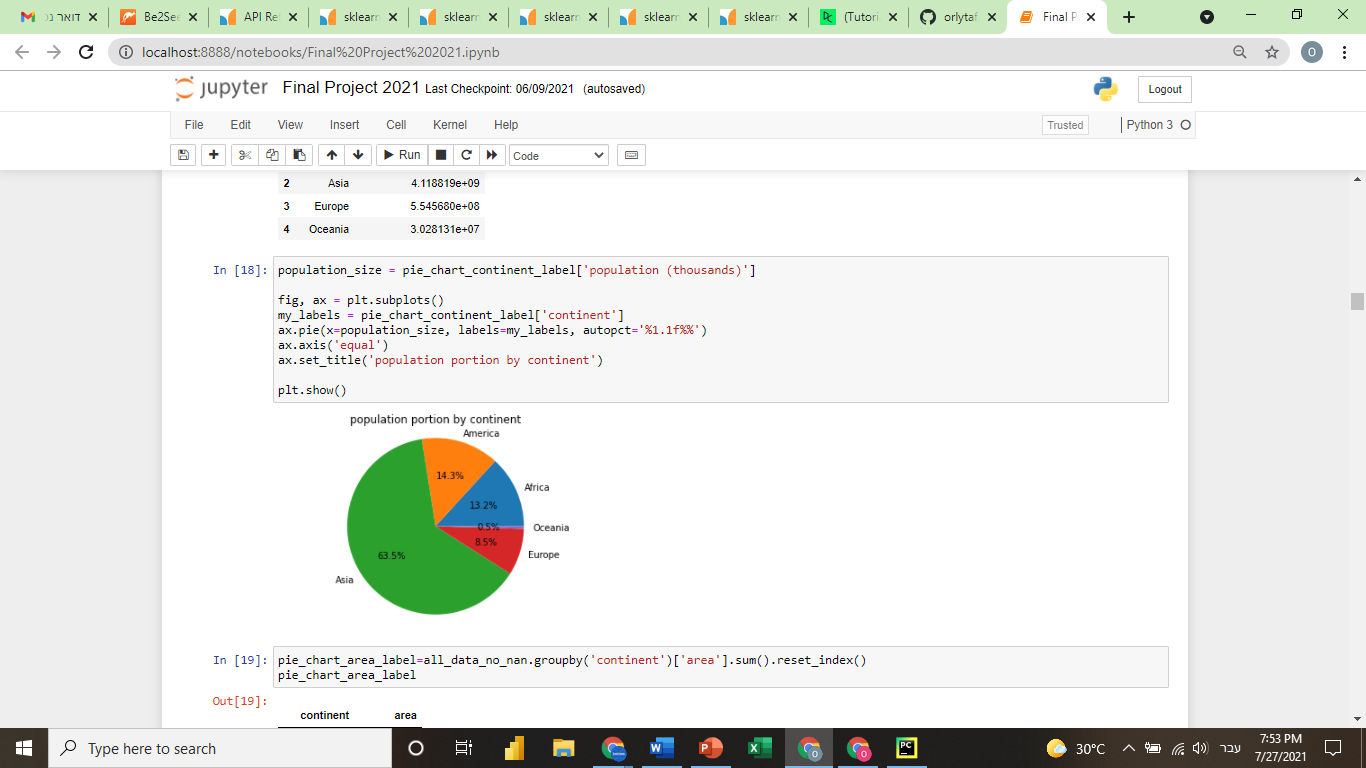
**מסקנות מהגרף:** זהו ***תרשים פיזור*** שבו נמדד מדד האושר ביחס לשנה שבה קיבלו המדינות את עצמאותן ( פרמטר שהוספתי בצורה עצמאית לדאטה פריים שלי). כאן ציפיתי שיתקיים קשר בין 2 הפרמטרים האלו וחשבתי שאמצא שככל ששנת העצמאות של המדינה תהיה מוקדמת יותר (משמע מדינה ותיקה יותר), כך מדד האושר יהיה גבוה יותר. אמנם הנתונים החריגים מאשרים את מה ששיערתי, אולם בפועל התרשים לא מעיד על קשר וודאי בין 2 הפרמטרים האלו.



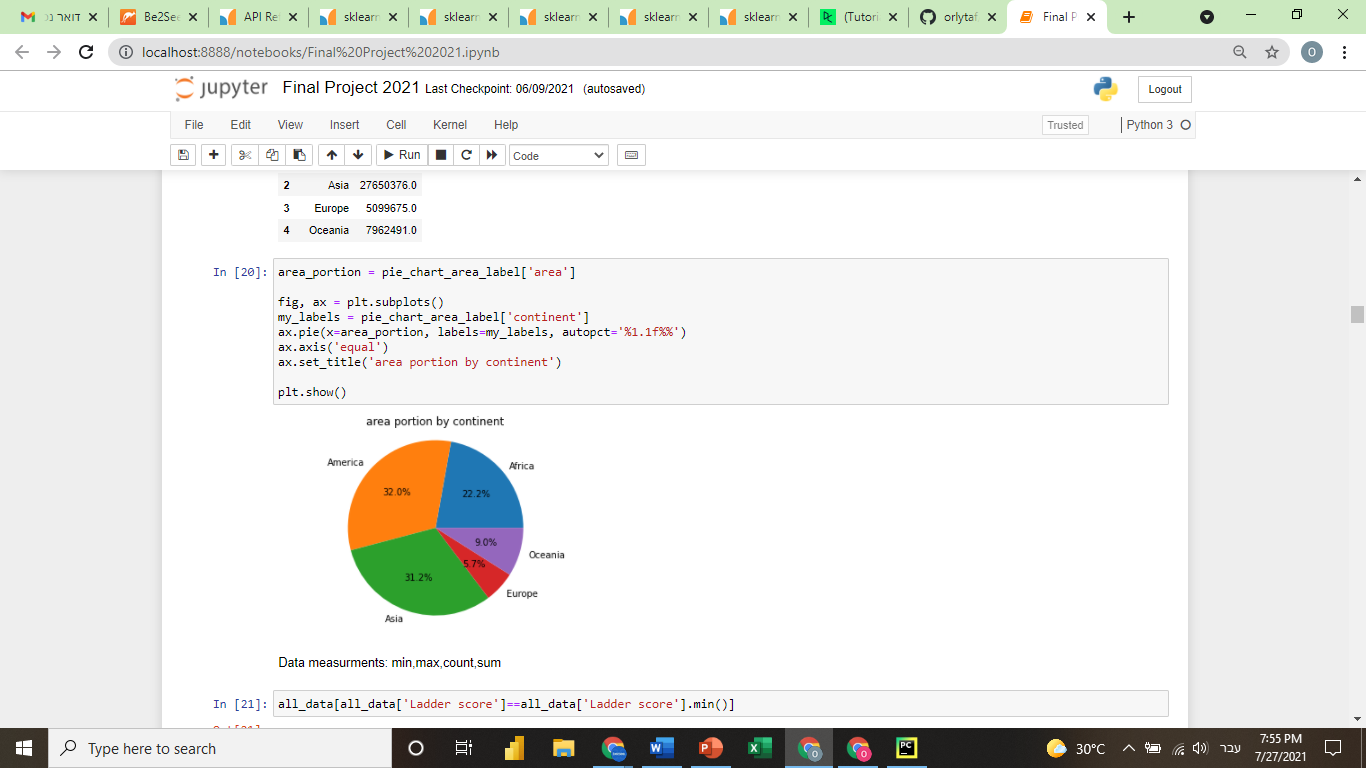
**מסקנות מהגרף:** זהו ***תרשים עוגה*** שבו סיווגתי את רמת החיים לפי קטגוריות: טוב,בינוני,נמוך ובדקתי מהו האחוז של המדינות שמשתייכות לכל קטגוריה. ניכר שרוב המדינות משתייכות לקטגוריה של רמת חיים טובה.  
  
**3. ויזואליזציות לפי יבשת:**



**מסקנות מהגרף:** ה-GDP הוא מעין מדד לעושר כלכלי של מדינה. ניתן לראות שה-GDP הגבוהים ביותר משויכים ליבשת אוקייניה (אוסטרליה וניו זילנד כלולות בתוכה) וליבשת אירופה, בעוד שהמדדים הנמוכים ביותר נצפו עבור יבשת אפריקה.



**מסקנות מהגרף:** היבשת שבה נמצא גודל האוכלוסיה הגבוה ביותר הינה אסיה, בעוד שבאוקייניה גודל האוכלוסיה הוא הנמוך ביותר.



**מסקנות מהגרף:** היבשת שגודל שטחה הוא הגבוה ביותר הינה אמריקה, בעוד שאירופה היא בעלת גודל השטח הנמוך ביותר.

**Neg\_mean\_squared\_error-** מודול ה-**sklearn.metrics** מאפשר יישום של פונקציות שעוסקות בטעויות של חיזויים שקורים מסיבה מסוימת. זו אחת מהפונקציות שקשורות למודל הרגרסיה.  
  
**Splitter-** פרמטר שמשויך למודל ML של Decision Tree. הפרמטר מקבל רשימה שמכילה פיצול קלטים בבדיקת test-train.  
  
**criterion(‘friedman\_mse’)-** פונקציה שמודדת את איכות הפיצול של הקלטים. friedman\_mse הוא בעצם שיפור של מדד ה-MSE הרגיל.  
  
**weights(‘distance’)-** פונקציית משקל שמשמשת ליצירת תחזיות. כאשר הפונקצייה מקבלת את הערך distance נקודות משקל נקבעות לפי ההופכי של המרחק שלהם. במקרה זה לשכנים קרובים יותר של נקודת שאילתה תהיה השפעה גדולה יותר מאשר לשכנים רחוקים יותר.