



Home Work

Data Science & Machine Learning

Data Analysis

NUMBER	DEADLINE	TOPICS
10	Tuesday, 1402/11/24 23:59:59 PM	Data Analysis, Pandas, Data Visualization

۱. سرگروه‌های هر گروه پاسخ تمرین را در کارپوشه، ارسال و پشتیبان فنی @Daneshkar_mentor را منشن کنید
۲. جواب سوالها را در یک فایل فشرده (شامل کدها و گزارش پروژه) جمع و ارسال کنید.
۳. در صورتی که سوال و یا ابهامی دارید در گروه چت تلگرامی بپرسید.
۴. لطفا نام فایل ارسالی خود را به صورت Group_{group_number}_Project1 با شماره گروه خود قرار دهید.



شرح پروژه:

از مجموعه داده tennis آنالیزهای زیر را استخراج کنید و توضیحات و نتیجه‌گیری‌های هر یک را در فایل گزارش ارائه دهید. ([شمای جداول پایگاه داده تنیس](#))

1. How many tennis players are included in the dataset?
2. What is the average age of the players?
3. Which player has the highest number of wins?
4. What is the longest match recorded in terms of duration?
5. How many sets are typically played in a tennis match?
6. Which country has produced the most successful tennis players?
7. What is the average number of aces per match?
8. Is there a difference in the number of double faults based on gender?
9. Which player has won the most tournaments in a single month?
10. Is there a correlation between a player's height and their ranking?
11. What is the average duration of matches?
12. What is the average number of games per set in men's matches compared to women's matches?
13. What is the distribution of left-handed versus right-handed players?
14. What is the most common type of surface used in tournaments?
15. How many distinct countries are represented in the dataset?
16. Which player has the highest winning percentage against top 10 ranked opponents?
17. What is the average number of breaks of serve per match?

علاوه بر تحلیل‌های فوق با خلاقیت خود ۷ تحلیل اضافه برای مجموعه داده تنیس اضافه و گزارش نمایید.

نکات:

۱. پروژه فوق به صورت گروهی قابل انجام است و تقسیم وظایف به صورت دلخواه و توافقی میان سرگروه و اعضای گروه صورت می‌گیرد (تقسیم وظایف میان تمام اعضای گروه باید صورت بگیرد).
۲. در صورت نبود امکان استخراج تحلیل‌های خواسته شده در بالا، می‌توانید آن تحلیل را نادیده بگیرید.
۳. داشتن ساختار درست و markdown در فایل(های) نوتبوک پروژه دارای نمره امتیازی می‌باشد.
۴. در کنار فایل کدها یک فایل **گزارش پروژه** نیز باید ارائه گردد که برای فرمت آن می‌توانید به انتهای این فایل رجوع کنید. برای قالب پروژه از فایل ضمیمه شده در این تمرین استفاده کنید.

ساختار فایل گزارش پروژه:

مجموعه داده:

در این بخش به صورت مختصر دانسته‌های خود را از مجموعه داده‌ها، جدوال و روال خواندن مجموعه داده‌ها را گزارش کنید. در نظر داشته باشید که نیازی به بیان مقدمات و بدیهیات مجموعه داده وجود ندارد و صرفاً کارهایی که در راستای شناخت مجموعه داده انجام شده است را گزارش کنید.

تحلیل و نتیجه گیری‌ها:

برای هر یک از تحلیل‌های فوق، بدون درج کردن کدهای موردنظر آن، نتایج و خروجی‌های خود را گزارش کنید. تحلیل‌های اضافه را به طور کامل در انتهای این بخش شرح و نتایج آن را گزارش کنید.

شرح وظایف:

به صورت مشخص وظایف هر یک از اعضای گروه در طول پروژه را در این بخش ذکر کنید.

پ.ن.: در صورت نیاز می‌توانید بخش‌های دیگری نیز به ساختار فوق اضافه کنید.

Happy Analyzing 😊