

بنام خدا

تمرین سوم

درس داده کاوی

- به زبان برنامه نویسی پایتون برنامه ای بنویسید که موارد خواسته شده زیر را انجام دهد.
- موارد خواسته شده را به دقت انجام دهید.
- از خروجی های بدست آمده در هر مرحله عکس بگیرید و در فایل پاسخ قرار دهید. پاسخ های خود را به طور کامل به صورت صوتی شرح دهید.
- فایل نهایی توضیحات به همراه کدها و خروجی های بدست آمده را به صورت یک فایل فشرده (rar یا zip) در لینک زیر آپلود کنید. اگر ویدیو قرار داده می شود بایستی حجم آن تا حد امکان کاهش داده شود.
<https://cloud.vru.ac.ir/s/KwoFWEECD7rz5CY>
- به کسی که بهترین و کاملترین محتوا برای هر تمرین را تولید کند (هم از نظر صحت پاسخ و هم از نظر گرافیکی و هم ارزش آموزشی برای دیگران) نمره اضافه نیز در نظر گرفته خواهد شد.

مصورسازی داده ها

1. فایل social_network_ads.csv را که به همراه تمرین قبل (سری دوم) ارسال شده است را خوانده و سن افراد (age) را با یک نمودار Box plot نمایش دهید.
2. فایل Mall_Customers را که به همراه تمرین قبل (سری دوم) ارسال شده است را خوانده و با استفاده از نمودار Joint plot میزان درآمد مردان و زنان را با یکدیگر مقایسه کنید.

استخراج قوانین انجمنی

کدهای ارائه شده در لینک زیر را پیاده سازی کرده و با استفاده از آنها برای مجموعه داده زیر موارد زیر را بدست آورید:

https://github.com/chonyy/apriori_python

مجموعه داده :

```
dataset = [['Milk', 'Onion', 'Nutmeg', 'Kidney Beans', 'Eggs', 'Yogurt'],  
           ['Dill', 'Onion', 'Nutmeg', 'Kidney Beans', 'Eggs', 'Yogurt'],  
           ['Milk', 'Apple', 'Kidney Beans', 'Eggs'],  
           ['Milk', 'Unicorn', 'Corn', 'Kidney Beans', 'Yogurt'],  
           ['Corn', 'Onion', 'Onion', 'Kidney Beans', 'Ice cream', 'Eggs']]
```

3. مجموع آیتم های پرتکرار با minsup=0.6
4. قوانین انجمنی با minsup=0.6 و minconf=0.7

موفق باشید