**مینی پروژه اول : Google Play Apps**

در ابتدا مراحل import کتاب خانه ها و پکیج های مورد نیاز و پس از آن load کردن دیتا را انجام میدهیم.

(importing the libraries & Load and Prepare Data)

در مرحلهی EDA پس از آشنایی کلی با دیتا و نوع داده های ستون ها، تمرکز روی آشنایی بیشتر با داده های ستون ها برای استفاده از آن ها در مراحل بعدی میرود.

در بخش Data Preprocessing، اقدامات لازم برای برطرف کردن nan ها انجام می شود و پس از آن با توجه به ستون ها، اقدام به عددی کردن ستون های Installs و Size برای استفاده ی احتمالی از آن در مرحله ی visualization نموده ام.

اکنون پس از EDA و Preprocessing امکان استفاده از داده ها برای مصور سازی های مورد نظر فراهم است که می توان نتایج زیادی را از این تصویر ها استخراج کرد که به چند مورد می پردازیم:

به ترتیب نمودارها:

#1 : دسته بندی های Family و Game و Tools بیشترین فراوانی را در بین category ها دارند و توزیع سایر دسته ها نیز قابل مشاهده است .

#2 : تعداد app های Free بسیار بیشتر از نوع Paid است که در نتیجه باعث بیشتر بودن تعداد نصب آنها نسبت به paid هاست.

هرچند که در مجموع app های paid وضعیت rating بهتری نسبت به نوع Free دارند.

#3 : در ادامه فراوانی app ها بر اساس content rating نمایش داده شده که نوزیع آنها مشخص است و ترتیب install آنها هم با همان نسبت فراوانیشان همخوانی دارد.

اینها تنها چند مورد از اطلاعات قابل مشاهده در نمودار ها و داده های ماست و با توجه به نوع هدف ما و دیدگاه ما به داده، میتوان خیلی جزئی تر به آن نگاه کرد و نتایج مختلفی را نیز از آن استخراج کنیم...