



پروژه پایانی

دومین دوره آموزش سریع پایتون با رویکرد علم داده

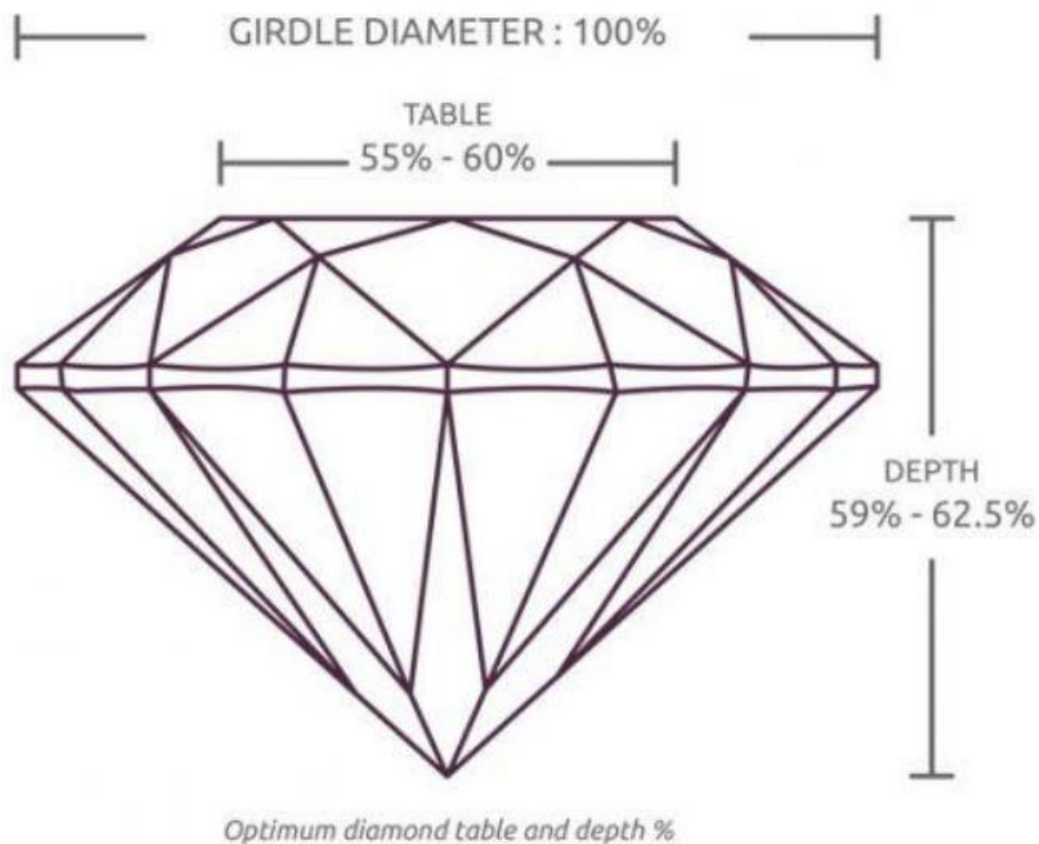
قسمت اول – الماس

در قسمت اول پروژه مجموعه داده ای مربوط به الماس های زینتی در اختیار شما قرار داده شده است. هدف از این قسمت، انجام فرآیند اکتشافی و تحلیلی در راستای بررسی تاثیر خصوصیات مختلف قطعات الماس بر قیمت گذاری آن می باشد.

خصوصیات 53,940 قطعه الماس در قالب یک فایل (diamonds.csv) در اختیار شما قرار داده شده است. همچنین قیمت هر قطعه الماس در قالب یک فایل دیگر (diamond_prices.csv) موجود است. در ادامه به معرفی ستون های موجود در داده گان می پردازیم.

ستون های diamonds.csv:

- **Id:** یک مقدار منحصر به فرد، شناسه الماس
- **carat:** قیراط
- **cut:** کیفیت برش، مقادیر به ترتیب افزایش کیفیت:
 - Fair (بدترین)
 - Good
 - Very Good
 - Premium
 - Ideal
- **color:** کیفیت رنگ، مقادیر به ترتیب افزایش کیفیت: حروف الفبای انگلیسی از D (بهترین) تا L (بدترین)
- **clarity:** میزان شفافیت الماس، مقادیر به ترتیب افزایش شفافیت:
 - I1 (بدترین)
 - SI2
 - SI1
 - VS2
 - VS1
 - VVS2
 - VVS1
 - IF (بهترین)
- **x:** طول (mm)
- **y:** عرض (mm)
- **z:** ارتفاع (mm)
- **depth:** $z/\text{mean}(x, y)$
- **table:** نسبت شعاع پهن ترین ناحیه بالایی الماس به شعاع کل



ستون های diamond_prices.csv:

- id: شناسه الماس
- price: قیمت

مراحل پروژه و خروجی های مورد انتظار بصورت جزئی در Notebook ای که در اختیارتان قرار گرفته است به صورت جزئی شرح داده شده اند. از شما انتظار می رود که توانایی تحلیل نتایج و ارائه آن ها را داشته باشید.

قسمت دوم – کرونا

در این قسمت یک داده سری زمانی مربوط به آمار کرونا در اختیار شما قرار گرفته است. داده ها شامل آمار مربوط به ابتال، مرگ و میر، و بهبود یافته ها (به صورت جهانی) می شود. مراحل و خروجی های مورد انتظار این قسمت نیز به صورت جزئی در Notebook ذکر شده اند.

نکات تکمیلی

- برای مواردی که در Notebook از شما توضیح یا تحلیل خواسته شده، توضیح خود را در یک سلول در Notebook قرار دهید.
- تسلط بر کد ها، و داشتن توانایی تحلیل خروجی ها هنگام ارائه پروژه الزامی است.
- محدودیتی در مورد استفاده از Library ها وجود ندارد؛ استفاده از کتابخانه های پایتون ای که در کلاس تدریس نشده اند بلامانع است. (در صورت تمایل)