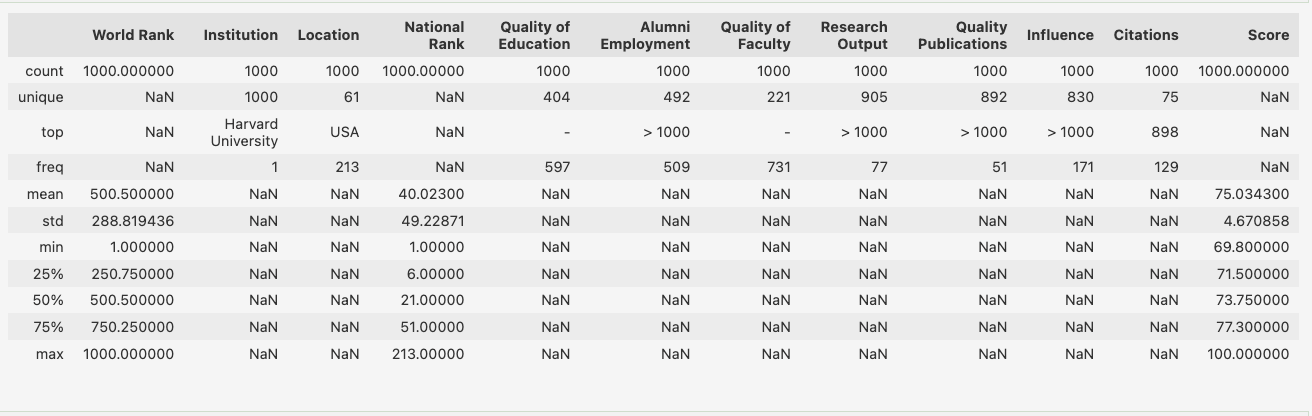
دیتاست1000×12 زیر برای انالیزبه شرح زیر است:



در نگاه اول با انالیز اولیه مقدار missing values را در جدول زیر مشاهده نمیکنیم زیرا missing values ها در ان به درستی درج نشده و بجای NaN از خط تیره(-) استفاده شده همچنین بعضی از مقدار ها به صورت > 1000 درج شده که کار انالیز را با مشکل روبرو میکند:



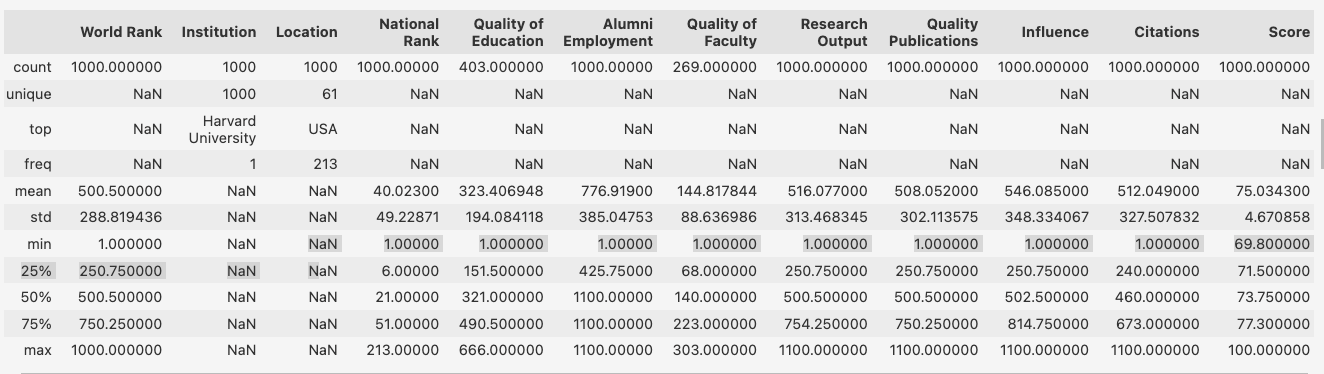
ابتدا > 1000 را تبدیل به 1100 شد برای اینکه در پلات ها قابل مشاهده و واضح باشد:



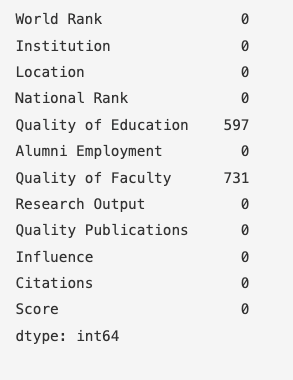
در مرحله بعدی لاین یا خط تیره(-) را تبدیل به NaN کردیم تا مقدار missing values برای هر ستون مشخص شود:



حال با انالیز کلی دیتاست مشخص میشود مقدار missing values ها چقدر است و ستون مربوط به ان کدام است:

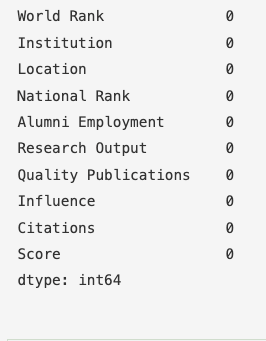


مشاهده میشود در ستون های Quality of Educationو Quaity of Faculty به ترتیب 597 و731 missing values,وجود دارد.



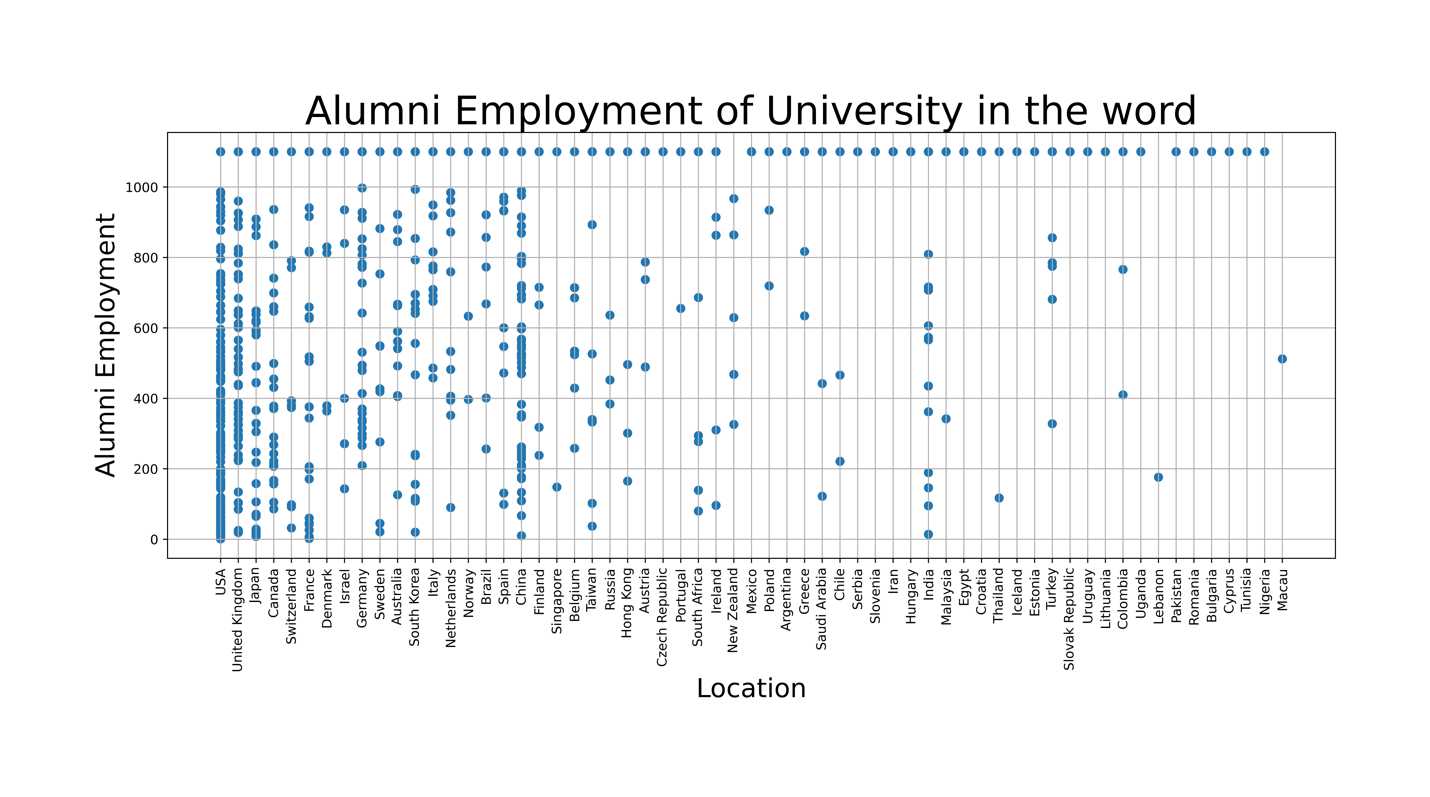
این missing values ها از دیتاست حذف شدند به علت تعداد قابل توجهشان در دیتاست و دیتاست جدید تبدیل 1000×10 شده است:



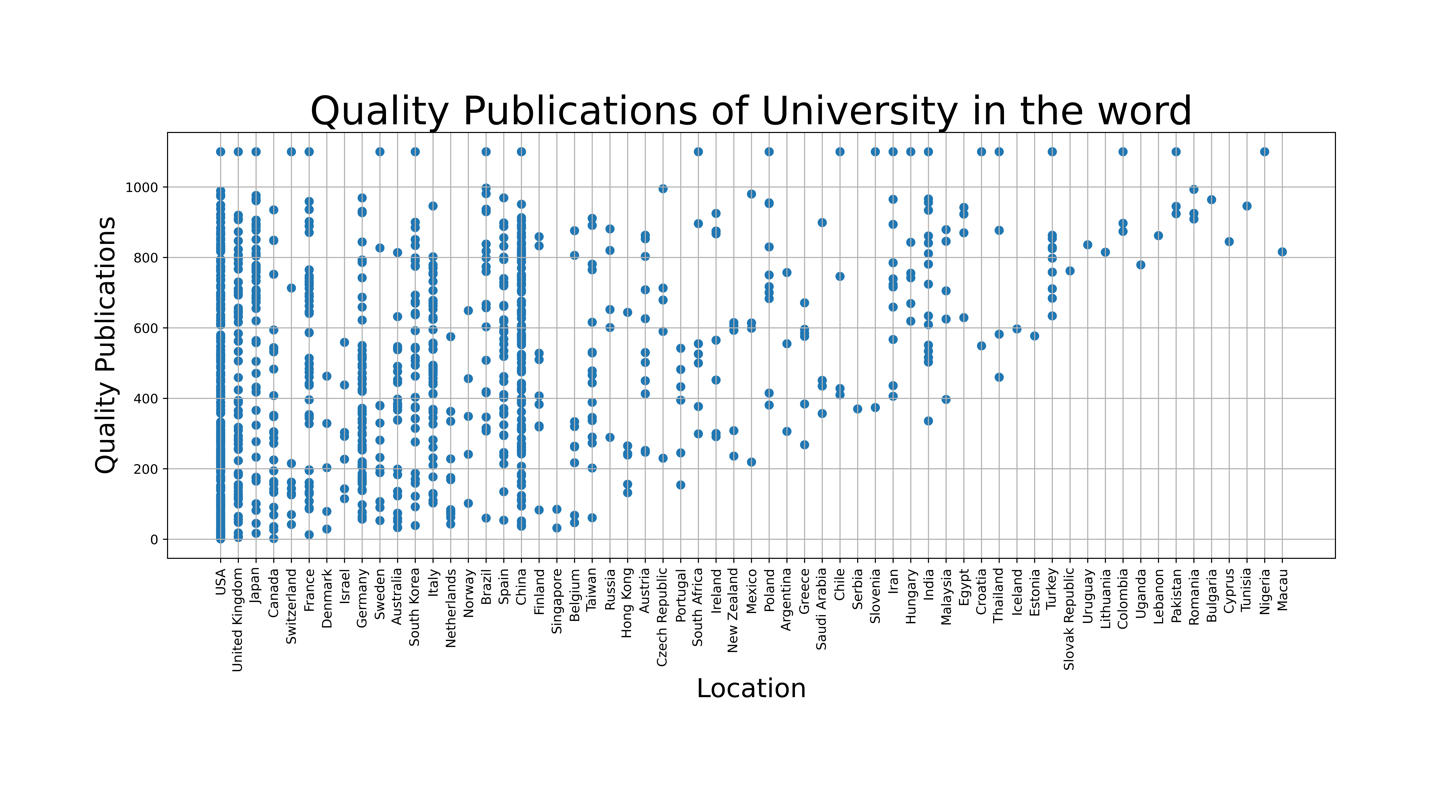


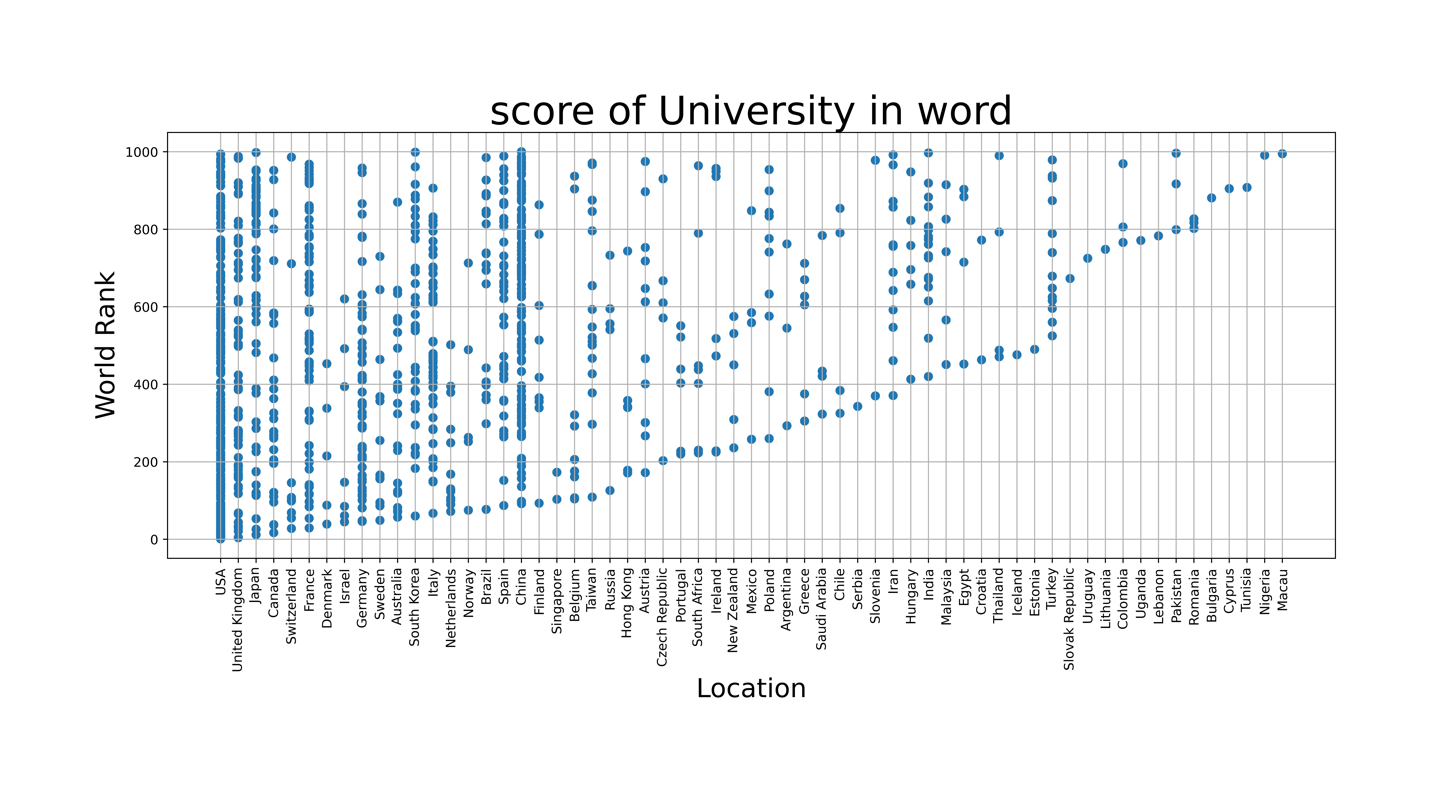
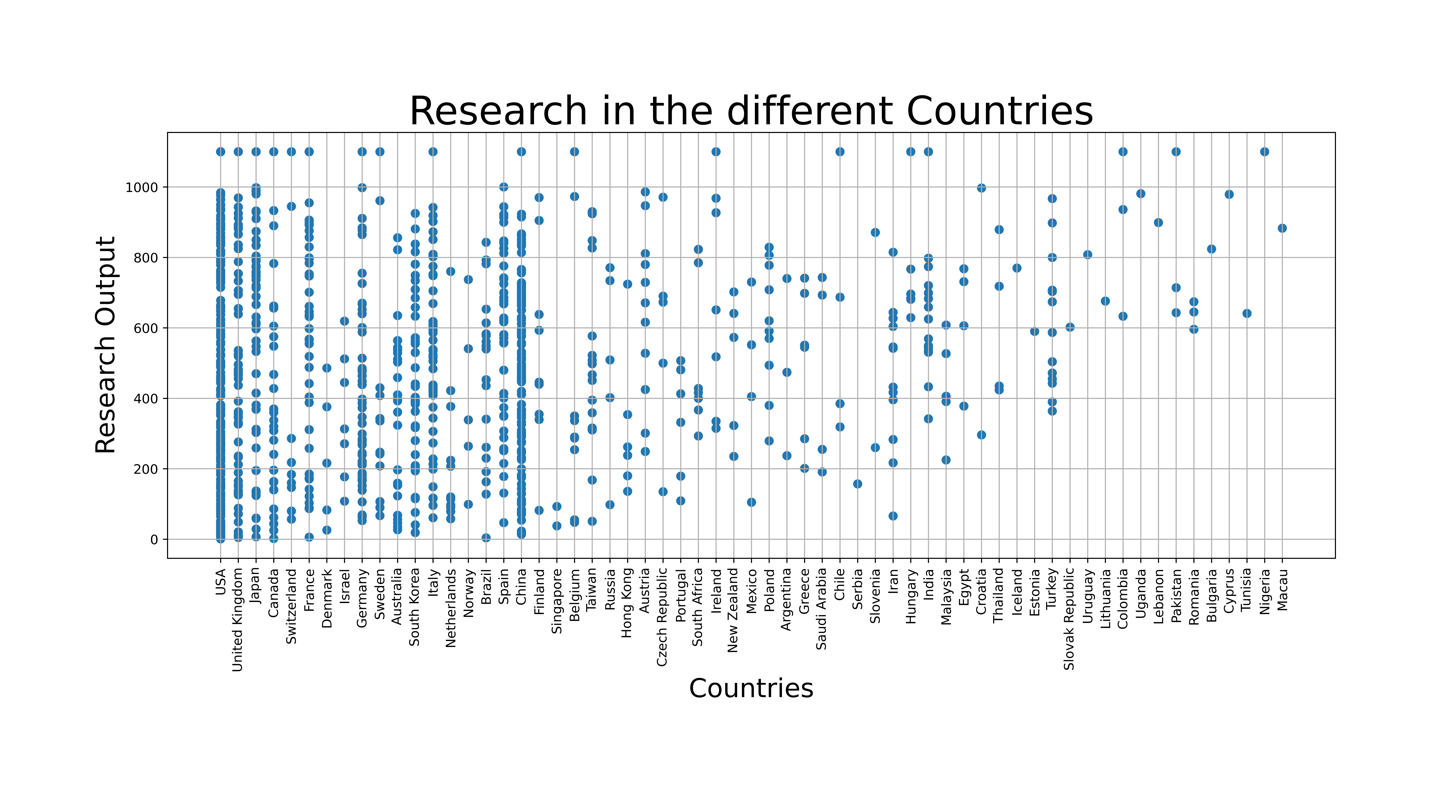
در پلات زیر که نشاندهنده ی میزان استخدام فارغالتحصیلان دانشگاهای مختلف است مشاهده میکنیم امریکا طیف وسیعی

از امتیاز ها را به خود اختصاص داده ( از۱ تا بالای ۱۰۰۰ ) و تعداد دانشگاهای مورد بررسی از چپ به راست از امریکا به ماکاءو کاهش یافته و به یک عدد رسیده که استخدام فارغالتحیصالن در ان مقدار تقریبا ۵۰۰ رو به خود اختصاص داده است

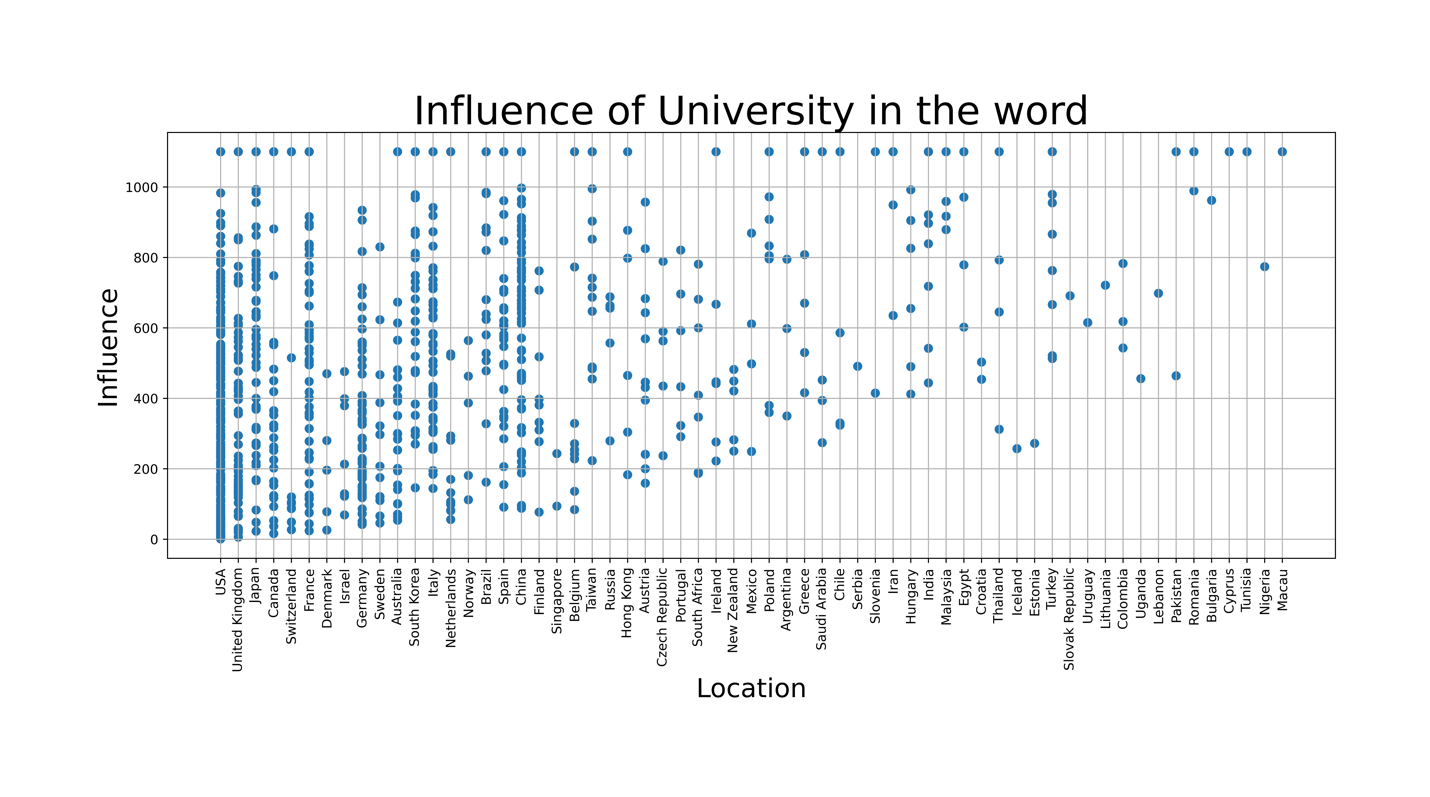


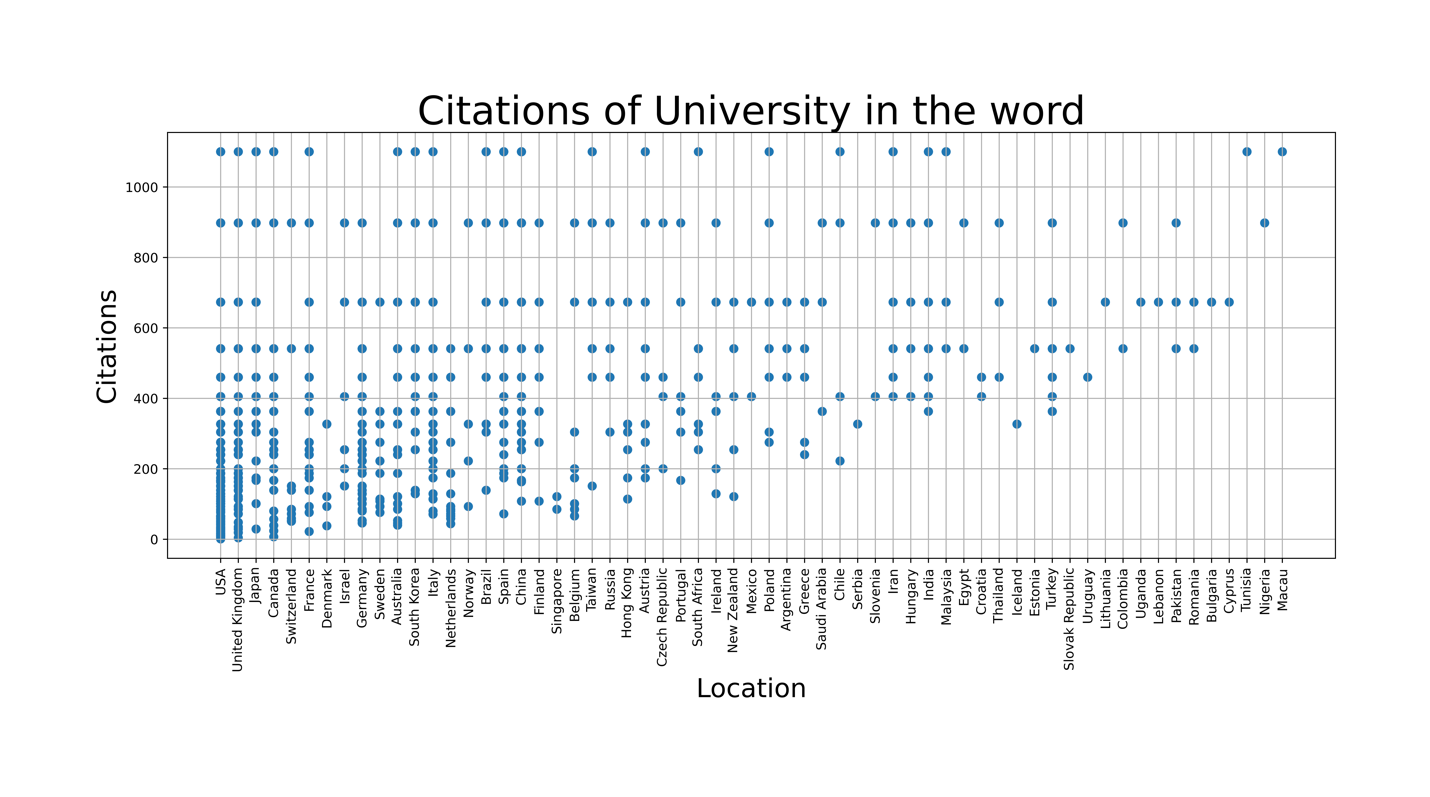
ااااااااااااااا

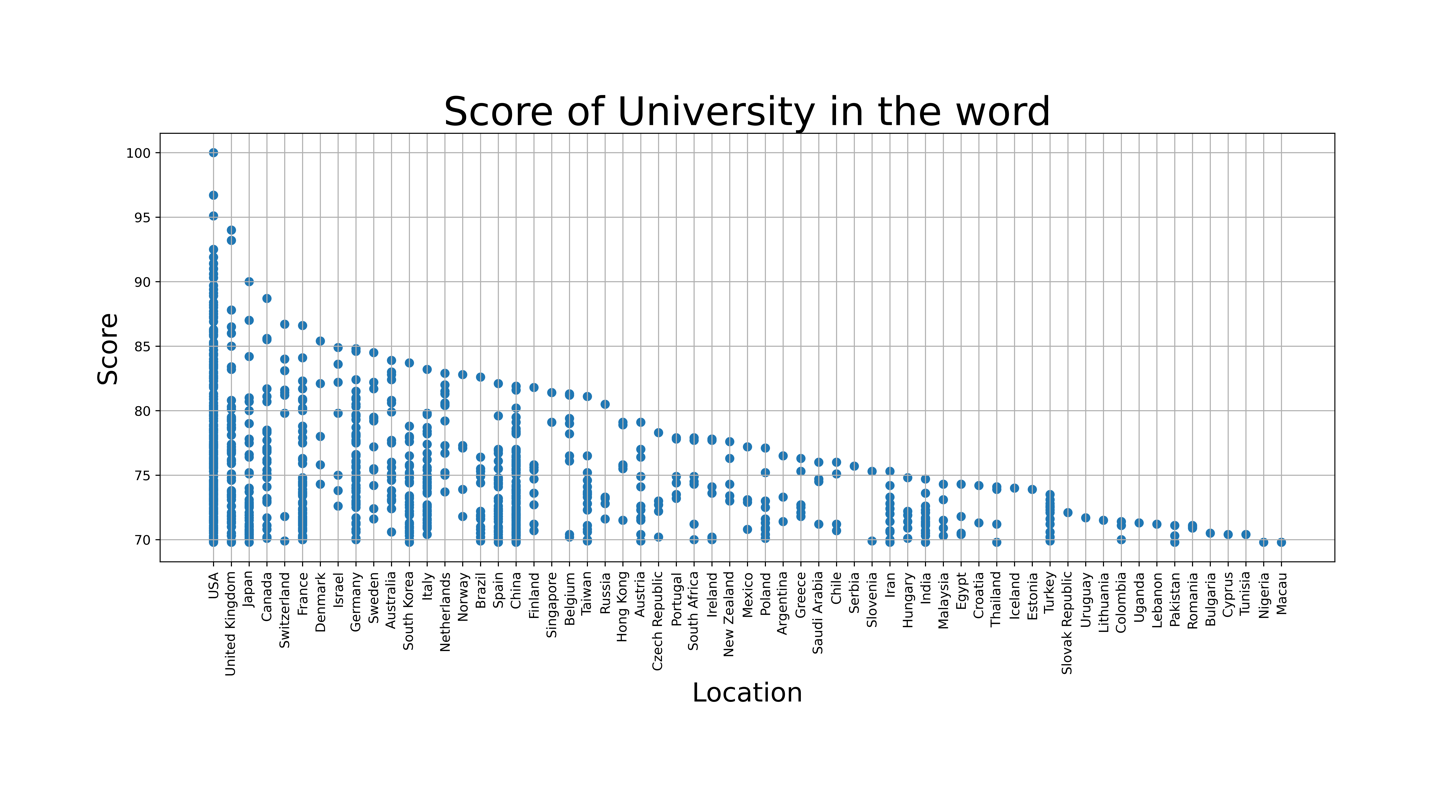


ببببببببببببپپپپپپپپپپپپ

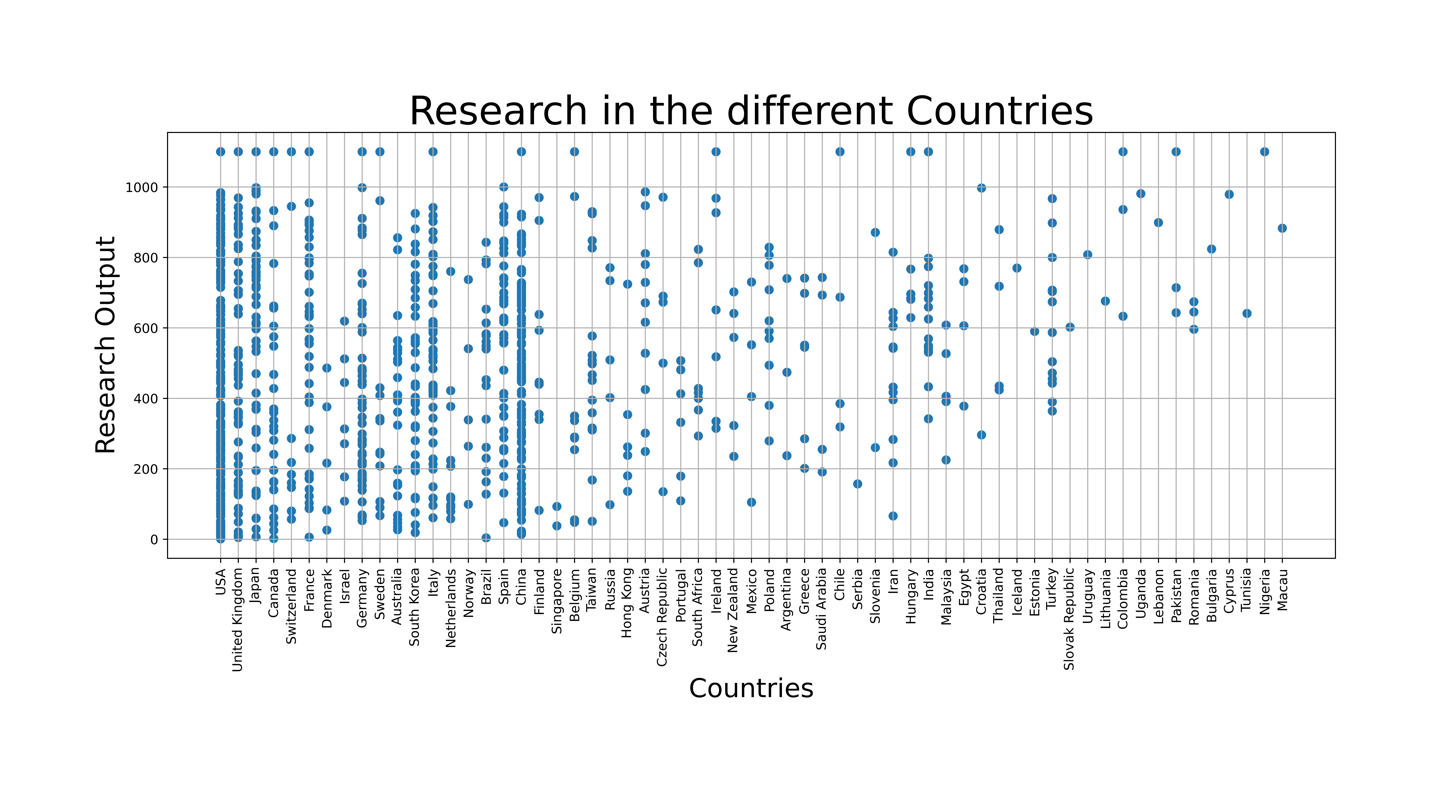
اتتتتتتتتتتتت

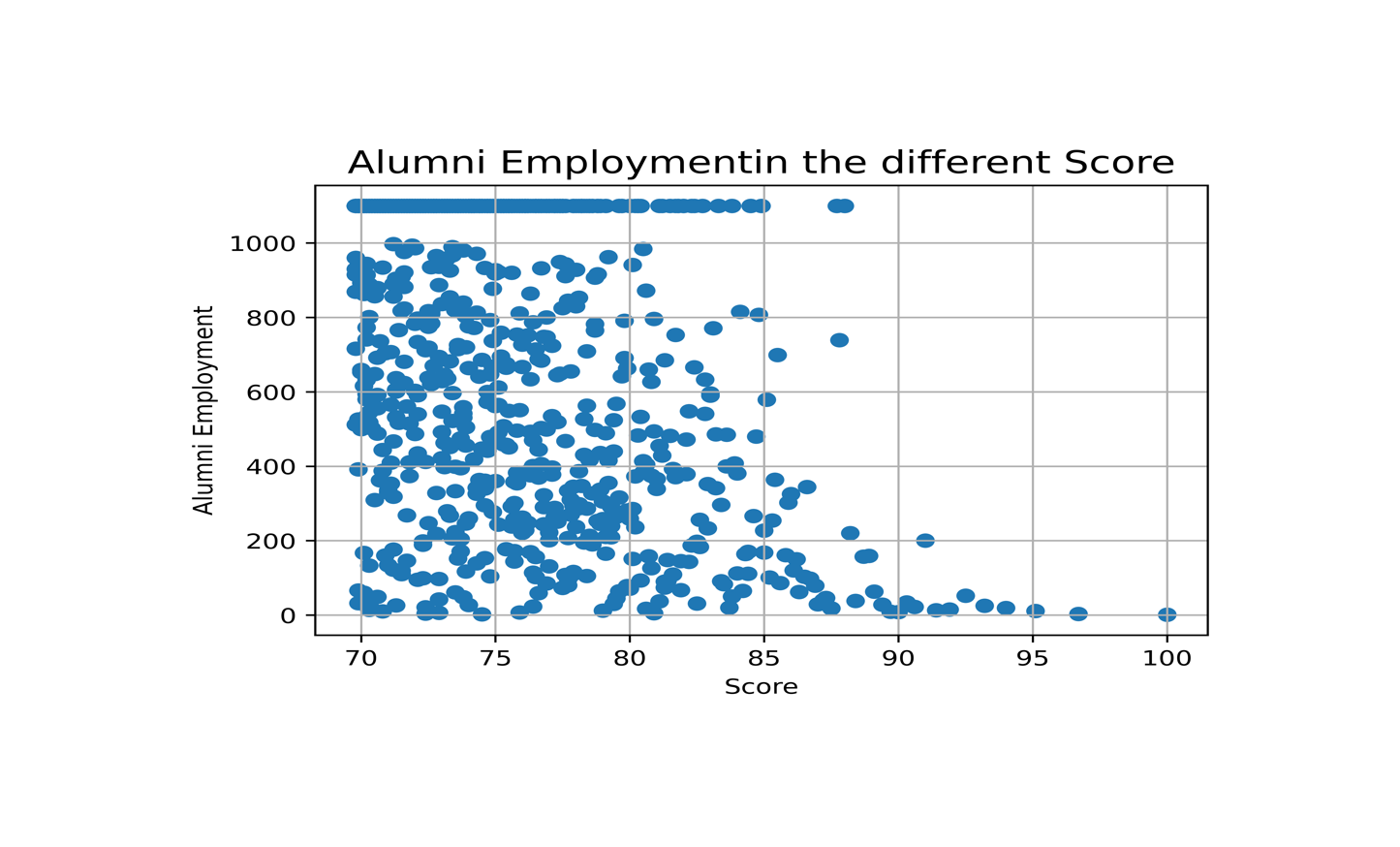


ثثثثثثثثثثثثثثثث

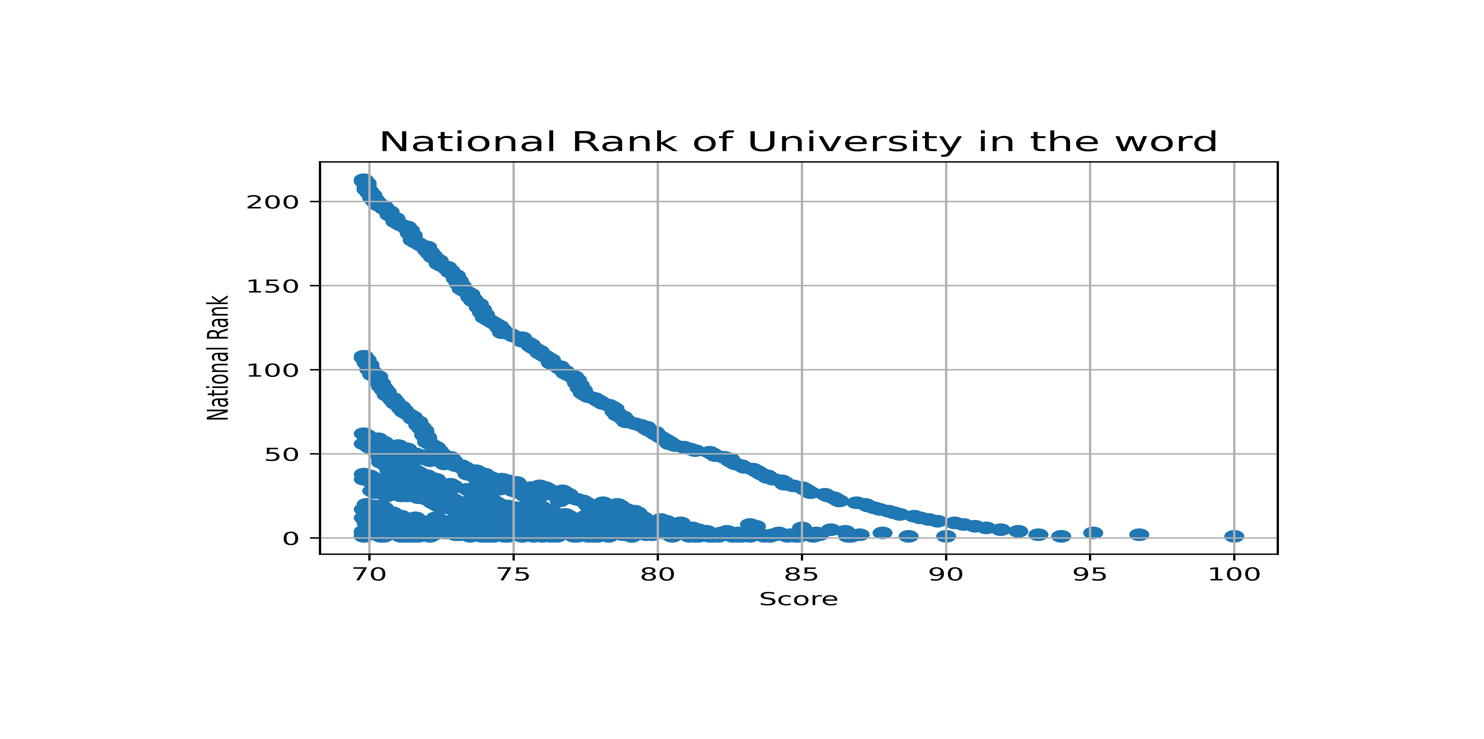
ییییییییییییی

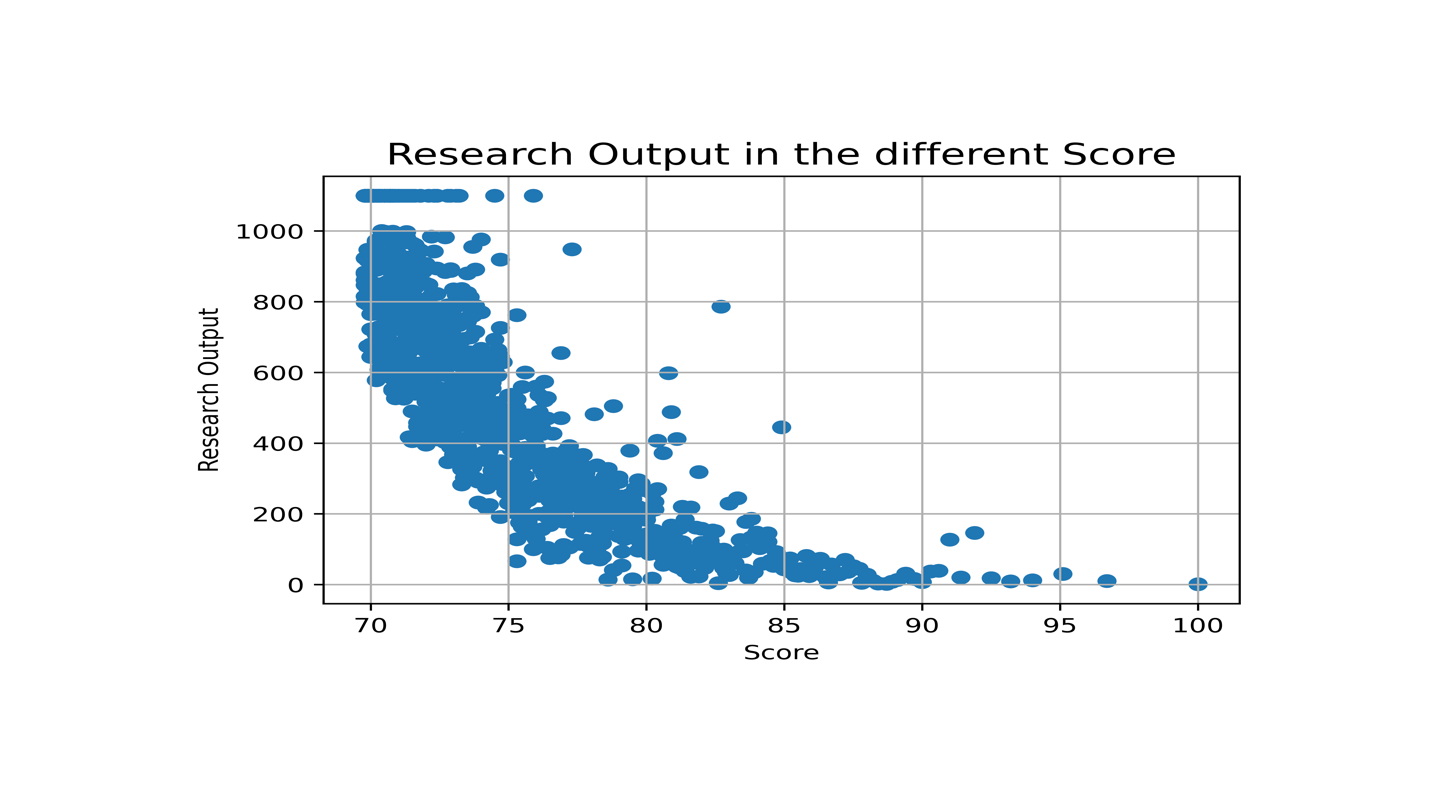
لللللللللللل



شششششششششششش

خخخخخخخخخخ



تتمتمتمتمتممتنننننن

