

چالش ورودی بوت کمپ

مهندسی داده

DATA ENGINEERING BOOTCAMP



چالش ورودی بوت کمپ مهندسی داده

از اینکه علاقمند به شرکت در بوت کمپ "مهندسی داده" رهنما کالج هستید، خوشحالیم. در فرآیند ثبت نام علاوه بر سوالاتی که در مورد رزومه و انگیزه شما پرسیده شده، لازم است به سوالات زیر پاسخ دهید و پاسخ تان را در فرم ثبت نام برای ما بارگذاری کنید.

سوال یک:

فرض کنید شما به تازگی به عنوان کارآموز Data Engineering در شرکت مریخیان مقیم زمین مشغول به کار شده است. از شما تقاضا شده مجموعه از داده ها را در یک DataBase رابطه ای به نام PostgreSQL وارد کنید. مدیر شما خواسته است تسک را با بخش کوچکی از دیتا در محیط لوکال تست کند. اکنون باید مراحل زیر را باهم طی کنیم.

۱- به کمک Docker یک PostgreSQL در محیط لوکال بالا بیاورید.

۲- یک Database با نام دلخواه بسازید و Table های با نام Customers و Orders با توجه به داده ها با تایپ مناسب ایجاد کنید.

۳- سپس داده ها را به Database منتقل کنید و Performance سیستم را رصد کنید. تلاش کنید بررسی کنید اگر حجم داده ورودی ۱۰، ۱۰۰ و ۱۰۰۰ برابر شود باز روشی که انتخاب کرده اید امکان پذیر است؟ راه های جایگزین چیست؟ (داده ها در فایل zip موجود است)

۴- یک کوئری بنویسید که نام هر مشتری و تعداد کل سفارشات که ثبت کرده اند را نمایش دهد.

۵- مجموع مبلغی که مشتریان در آمریکا (USA) خرج کرده اند را به تفکیک هر مشتری بازیابی کنید.

۶- سه مشتری برتر که بیشترین مقدار را پرداخت کرده اند را به ترتیب نزولی از لحاظ مجموع مبلغ خرج شده بازیابی کنید.

۷- یک کوئری بنویسید که هر کشور را به همراه تعداد مشتریان و مجموع مبلغ تمامی سفارش ها از آن کشور فهرست کند. تمامی کوئری ها را در فایل جدا قرار دهید.

سوال دو:

اکنون باید یک Service بنویسید تا هر دقیقه خروجی کوئری-۷ سوال یک را از db خوانده و در Redis ذخیره سازی کند. آیا راه بهینه تری می توانید معرفی کنید؟ ترجیحا از زبان Java یا حداکثر از Python برای پیاده سازی کنید. لازم است برای کار به کمک Docker-Compose یک Redis و PostgreSQL بالا بیاورید.

دقت کنید!

- لطفا هرچا احساس کردید سوالات ابهام دارند فرض مطلوب خود را بنویسید و حل سوالات ادامه دهید.
- لطفا از حل سوالات به کمک ابزار های هوش مصنوعی امتناع کنید تا ارزیابی شایسته تری صورت گیرد.
- پاسخ های خود را باید در قالب یک فایل Zip ارسال کنید که نام آن ترکیبی از نام و نام خانوادگی شما به طور مثال، Ali.Akbari باشد.
- هرچند جست و جوی فردی بخشی از مراحل حل مسئله است با این وجود نهایتا ما در ارزیابی ها و بررسی پاسخ ها مبنا را بر صداقت شما می گذاریم و فرض می کنیم به این موضوع واقف هستید که پاسخ غیرصادقانه شما، ما را در ارزیابی صحیح دچار گمراهی می کند و ممکن است وارد دوره ای شوید که متناسب با نیازتان نیست.
- با توجه به محدود بودن ظرفیت دوره، چالش ها به ترتیب ارسال، بررسی و ارزیابی می شوند و با پر شدن ظرفیت، مهلت ارسال چالش به **اتمام** می رسد.

به امید دیدارتان هستیم