

بوت کمپ تحلیل داده

DATA ANALYSIS BOOTCAMP



چالش ورودی بوتکمپ تحلیل داده

از اینکه علاقمند به شرکت در بوت کمپ تحلیل داده رهنماکالج هستید، خوشحالیم. در فرآیند ثبت نام علاوه بر سؤالاتی که در مورد رزومه و انگیزه شما پرسیده شده، لازم است به سؤالات زیر نیز پاسخ دهید و جوابتان را در فرم ثبت نام برای ما آپلود کنید.

داده‌ای که در **این آدرس** در اختیار شما قرار گرفته حاوی اطلاعات مربوط به حدود یک میلیون آگهی فروش اتومبیل در آمریکا است که در سپتامبر ۲۰۲۰ از یک وبسایت خرید و فروش خودرو استخراج شده است. در نتیجه، داده‌ها مربوط به خودروهای فروش نرفته در تاریخ استخراج داده‌هاست.

برای تحلیل داده‌ها از هر ابزاری که در اختیار دارید می‌توانید استفاده کنید. اگر با زبان برنامه‌نویسی SQL آشنایی دارید، می‌توانید فایل داده را در **این آدرس** آپلود کنید و همان‌جا به صورت آنلاین کوئری بنویسید. توضیحات مربوط به ستون‌های فایل داده در ادامه آمده است.

۱. کدام شرکت سازنده خودرو دارای بیشترین تعداد آگهی است؟ در میان خودروهای این شرکت، کدام مدل و رنگ بیشتر آگهی شده است؟
۲. پراکندگی مدل‌های خودرو در سه شهر پراکندگی کدامند؟
۳. میانگین قیمت خودروهای آگهی شده چند دلار است؟
۴. شرکت‌های سازنده را بر حسب متوسط قیمت خودروهایشان مرتب کنید.
۵. فرض کنید خودروهای با حجم موتور کمتر از ۲۰۰۰ سی‌سی را کلاس A، خودروهای با حجم موتور بین ۲۰۰۰ سی‌سی تا ۴۰۰۰ سی‌سی را کلاس B و خودروهای با حجم موتور بالاتر از ۴۰۰۰ سی‌سی را کلاس C نام‌گذاری کنیم. در این صورت، هر کلاس چه سهمی از خودروهای ساخته شده توسط سه سازنده‌ی دارای بیشترین آگهی خودروهای با سیستم دنده‌ای اتومات را به تفکیک شامل می‌شود؟
۶. آیا حجم موتور تأثیری بر میزان مصرف سوخت دارد؟
۷. سهم خودروهای هیبریدی از کل خودروهای آگهی شده برای کدام سازنده بیشتر است؟
۸. چند درصد از خودروهای آگهی شده به عنوان خودروی صفر دسته بندی می‌شوند؟
۹. خودروهای کمتر از دو سال ساخت از لحاظ قیمتی چه تفاوتی با یکدیگر دارند؟
۱۰. چند درصد خودروها تصادفی‌اند؟
۱۱. خودروهای کدام برند به طور متوسط زودتر فروش می‌روند؟
۱۲. با توجه به آگهی‌ها، چه عواملی ممکن است باعث طولانی‌تر شدن فرآیند فروش خودرو شود؟
۱۳. ماه‌های سال را بر حسب بیشترین تعداد خودروی آگهی شده در آنها مرتب کنید. آیا آگهی شدن خودروها در ماه‌های سال از الگوی زمانی خاصی پیروی می‌کند؟
۱۴. سه شهری را که در آنها خودروهای قدیمی‌تری آگهی شده‌اند به ترتیب نام ببرید.
۱۵. بیشترین تعداد خودروی تصادفی به کدام سازنده تعلق دارد؟
۱۶. آیا می‌توان ادعا کرد در بین خودروهای آگهی شده، خودروهایی که در ۱۰ سال اخیر ساخته شده‌اند تعداد صندلی کمتری نسبت به خودروهای قدیمی‌تر دارند؟
۱۷. توزیع خودروها را از لحاظ نوع بدنه در پنج شهر پرجمعیت آمریکا (در پایان سال ۲۰۲۳) با یکدیگر مقایسه کنید.

دقت کنید

● پاسخ‌های خود را در قلب یک فایل pdf ارسال کنید که نام آن ترکیبی از نام و نام خانوادگی شما (به طور مثال، Ali.Akbari.pdf) باشد. برای شرکت در چالش لزومی به حل تمام سؤالات نیست، البته حل درست سؤالات بیشتر، احتمال قبولی شما را بالاتر می‌برد.

● هرچند جست‌وجوی فردی بخشی از مراحل حل مسئله است با این وجود نهایتاً ما در ارزیابی فرم و انجام تمرین‌ها مبنای بر صداقت شما می‌گذاریم و فرض می‌کنیم شما به این موضوع واقف هستید که پاسخ غیرصداقانه شما، ما را در ارزیابی صحیح دچار گمراهی می‌کند و ممکن است وارد دوره‌ای شوید که متناسب با نیازتان نیست.

نام ستون

توضیح

شناسه یکتای آگهی	listing_id
تاریخ انتشار آگهی	listed_date
شهر	city
شرکت سازنده	manufacturer
مدل	model
سال ساخت	year
رنگ	color
قیمت (دلار)	price
نوع بدنه	body_type
حجم موتور (سی سی)	engine_volume_cc
مصرف سوخت (کیلومتر قابل پیمایش به ازای هر لیتر سوخت)	fuel_consumption
نوع سوخت	fuel_type
کمتر از دو سال از تولید خودرو گذشته باشد 1	is_new
تعداد صندلی‌ها	maximum_seating
کیلومتر	mileage
قدرت موتور (اسب بخار)	horsepower
سیستم دنده (دستی: M و Dual Clutch - اتومات: A و CVT)	transmission
تصادفی 1	has_accidents
زمان طی شده از انتشار آگهی (روز)	days_on_market

منتظران هستیم.