

# نمونه سوال برنامه‌نویسی

دوره استادی هوش مصنوعی درسمن

تجزیه و تحلیل داده‌ها



موضوع: انجام فرایند ETL روی داده‌های فایل ضمیمه مربوط به محصولات یک فروشگاه

درجه سختی سوال: ☐ آسان ☒ متوسط ☐ سخت

تمرین شماره ۱:

پروژه ای را پیاده سازی کنید که درون آن عملیات Extraction, Transform و Load را روی داده‌های داخل فایل orders.csv (سفارش مشتریان فروشگاه) انجام دهد و در نهایت فایل حاصل از خروجی فرایند ETL را با نام target.csv ذخیره کند.  
(در فرایند پروژه می توانید از کتابخانه های `numpy`, `pandas`, `sklearn`, `plotly` و یا هر کتابخانه کاربری دیگری استفاده کنید)

عملیاتی که می توانید در پیاده سازی پروژه انجام دهید:

- خواندن فایل داده ای با پسوند CSV
- بررسی داده‌های Nan و حذف یا جایگزینی آنها
- تبدیل رشته‌های نویزی در ستون‌های عددی
- گسسته سازی داده ها
- رمزگذاری داده‌های دسته‌ای
- نرمال سازی و استاندارد سازی داده‌ها
- ذخیره سازی داده‌های نهایی