

1) sales_df_train.csv – данные по продажам за скользящий год для обучения.

Столбцы:

st_id – захэшированное id магазина;
pr_sku_id – захэшированное id товара;
date – дата;
pr_sales_type_id – флаг наличия промо;
pr_sales_in_units – число проданных товаров без признака промо;
pr_promo_sales_in_units – число проданных товаров с признаком промо;
pr_sales_in_rub – продажи без признака промо в РУБ;
pr_promo_sales_in_rub – продажи с признаком промо в РУБ;

2) pr_df.csv – данные по товарной иерархии.

От большего к меньшему pr_group_id - pr_cat_id - pr_subcat_id - pr_sku_id.

Столбцы:

pr_group_id – захэшированная группа товара;
pr_cat_id – захэшированная категория товара;
pr_subcat_id – захэшированная подкатегория товара;
pr_sku_id – захэшированное id товара;
pr_uom_id (маркер, обозначающий продаётся товар на вес или в ШТ).

3) pr_st.csv – данные по магазинам.

Столбцы:

st_id – захэшированное id магазина;
st_city_id – захэшированное id города;
st_division_code_id – захэшированное id дивизиона;
st_type_format_id – id формата магазина;
st_type_loc_id – id тип локации/окружения магазина;
st_type_size_id – id типа размера магазина;
st_is_active – флаг активного магазина на данный момент.

4) sales_submission.csv – пример файла с результатом работы модели прогноза спроса.

Необходимо подготовить файл в таком же формате, заполнив колонку target предсказаниями (по умолчанию колонка заполнена нулями).

Столбцы:

st_id – захэшированное id магазина;
pr_sku_id – захэшированное id товара;
date – дата (день);
target – спрос в ШТ.