



دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)

گزارش بخش اول
پروژه‌ی درس «پایگاه داده»

علی علی محمدی^س

۹۶۱۳۰۲۷

در بخش اول پروژه، تلاش بر آن بوده است که دیتاست‌های محصولات و تاریخچه‌ی خرید شرکت «دیجی‌کالا» را به حالت 5NF در پایگاه‌داده‌ی رابطه‌ای ذخیره شوند.

در ابتدا، ۶ دسته‌بندی محصولات، در نظر گرفته شدند.

با توجه به نحوه‌ی ذخیره شدن ویژگی‌های محصولات، لازم بود تا ویژگی‌های مشترک هرکدام از دسته‌بندی‌ها را استخراج کرده و جدول مربوط به آن دسته‌بندی را در پایگاه داده بسازیم.

توجه. برخی از ویژگی‌های محصولات، با این که در قسمت ویژگی‌های دیتاست آورده شده‌اند، اما به ازای هیچ‌کدام از محصولات آن دسته، مقدار ندارند. به عنوان مثال، تمام کیبوردها ویژگی رنگ را دارند اما هیچ‌کدام رنگشان مشخص نشده است.

برای صرفه‌جویی در استفاده از فضا، این‌گونه ویژگی‌ها در ساخت جدول‌ها لحاظ نشده‌اند و فقط ویژگی‌هایی که مقدار دارند، قرار داده شده‌اند.

توجه شود که برای ساخت جدول‌ها و ذخیره‌سازی اطلاعات از MySQL استفاده شده است.

سپس با استفاده از یک Script که به زبان Python نوشته شده است، فایل مربوط به دیتاست محصولات خوانده شده و با توجه به دسته‌بندی هر محصول، در صورت نیاز، مشخصات آن در پایگاه داده وارد می‌شود. (ساخت جدول‌ها نیز توسط همین Script انجام می‌شود)

همچنین این Script، داده‌های مربوط به تاریخچه‌ی خرید مشتریان را نیز در یک جدول جداگانه در پایگاه داده ذخیره می‌کند تا بتوان روی فیلد تاریخ آن، index قرار داد و کوئری‌های مرتبط به آن فیلد را تحلیل کرد.

سپس ۵ کوئری مرتبط به فیلد تاریخ، بر روی جدول تاریخچه‌ی خرید مشتریان اجرا می‌شوند و با استفاده از کلیدواژه‌ی EXPLAIN از نحوه‌ی اجرای این کوئری‌ها توسط DBMS مطلع خواهیم شد.

در صورتی که این کوئری‌ها از index که روی فیلد تاریخ قرار داده شده است، استفاده کنند، پیغام “Our index was USED!” به نمایش درخواهد آمد و در غیر این صورت، پیغام “Our index was NOT USED!” به نمایش درخواهد آمد.

توجه شود که نتایج مربوط به این بخش، ممکن است بر اساس سخت‌افزاری که این برنامه بر روی آن اجرا می‌شود متفاوت باشد.

حال به بررسی نتایج به دست آمده در سخت‌افزار اینجانب پرداخته می‌شود.

• کوئری اول:

```
EXPLAIN SELECT cart_finalize_datetime, COUNT(*) FROM Buying_History GROUP BY cart_finalize_datetime HAVING COUNT(*) > 1;
```

	id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
►	1	SIMPLE	Buying_History	HULL	index	date_index	date_index	6	HULL	199657	100.00	Using index

✓ از index استفاده شده است.

• کوئری دوم:

```
EXPLAIN SELECT AVG(order_gross_amount) FROM Buying_history WHERE item_quantity > 1 GROUP BY cart_finalize_datetime HAVING COUNT(*) > 2;
```

	id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
►	1	SIMPLE	Buying_history	HULL	index	date_index	date_index	6	HULL	199657	33.33	Using where

✓ از index استفاده شده است.

- کوئری سوم:

```
EXPLAIN SELECT DISTINCT cart_finalize_datetime FROM buying_history;
```

	id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
▶	1	SIMPLE	buying_history	<small>NULL</small>	index	date_index	date_index	6	<small>NULL</small>	199657	100.00	Using index

✓ از index استفاده شده است.

- کوئری چهارم:

```
EXPLAIN SELECT cart_finalize_datetime FROM buying_history ORDER BY cart_finalize_datetime ASC;
```

	id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
▶	1	SIMPLE	buying_history	<small>NULL</small>	index	<small>NULL</small>	date_index	6	<small>NULL</small>	199657	100.00	Using index

✓ از index استفاده شده است.

- کوئری پنجم:

```
EXPLAIN SELECT id, cart_finalize_datetime FROM Buying_History WHERE cart_finalize_datetime >= "2018-05-05";
```

	id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
▶	1	SIMPLE	Buying_History	<small>NULL</small>	range	date_index	date_index	6	<small>NULL</small>	88966	100.00	Using where; Using index

✓ از index استفاده شده است.