

**گروه کامپیوتر** – **دانشکده مهندسی** – **دانشگاه فردوسی مشهد**

**گزارش فاز اول پروژه**

**درس: مبانی داده کاوی**

**ارائه دهندگان:**

**سبحان مرادیان دقیق - فاطمه چاجی**

**استاد:**

**دکتر بهشید بهکمال**

**پاییز 1399**

# **تعریف قوانین معتبر بودن داده ها**

# **ارزیابی صحت داده ها**

* 1. **دیتاست محصولات**
* **ويژگي url\_code**

مقدار این ویژگی را با توجه به ویژگی product\_title\_fa بدست می‌آوریم. بدین صورت که در بین کلمات مقدار ویژگی – قرار می‌گیرد.

* **ویژگی category\_title\_fa**

مقدار این ویژگی زيرمجموعه مقدار ویژگی title\_alt در فیلد متناظر می‌باشد.

* **ویژگی category\_title\_en**

برای بررسی صحت مقدار وارد شده برای این ویژگی، می‌توان از مقدار وارد شده برای ویژگی دسته بندی فارسی محصول استفاده کرده و مطابقت واژه انگلیسی بکار رفته با واژه فارسی بکار رفته را بررسی کرد.

* 1. **دیتاست تاریخچه خرید مشتریان**
* **ویژگی city\_name\_fa**

برای بررسی صحت مقدار وارد شده برای این ویژگی، لیست تمامی شهر های ایران را با کمک [ویکی پدیا](https://fa.wikipedia.org/wiki/%D9%81%D9%87%D8%B1%D8%B3%D8%AA_%D8%B4%D9%87%D8%B1%D9%87%D8%A7%DB%8C_%D8%A7%DB%8C%D8%B1%D8%A7%D9%86) ذخیره می‌کنیم. نام وارد شده برای این ویژگی در صورتی صحیح خواهد بود که یکی از مقادیر موجود در این لیست باشد.

* 1. **دیتاست تاریخچه محصولات**
* **ویژگی هاي selling\_price و rrp\_price**

با مقایسه دو ویژگی selling\_price و rrp\_price و با توجه به تعریف rrp\_price[[1]](#footnote-1)، می‌توان نتیجه گرفت همواره رابطه زیر بین مقادیر وارد شده برای این ویژگی برقرار است:

selling\_price >= rrp\_price

* **ویژگی هاي buy\_priceو base\_price**

با توجه به تعريف دو ویژگی buy\_priceو base\_price می‌توان نتیجه گرفت همواره رابطه زیر بین مقادیر وارد شده برای این ویژگی برقرار است:

buy\_price <= base\_price

* **ویژگی هاي start\_atو created\_at**

با توجه به تعريف دو ویژگی start\_atو created\_at می‌توان نتیجه گرفت همواره رابطه زیر بین مقادیر وارد شده برای این ویژگی برقرار است:

* start\_at >= created\_at
  1. **دیتاست کیفیت محصولات**
  2. **دیتاست نظرات**

# **ارزیابی کامل بودن داده ها**

1. recommended retail price: the price that the company that makes a product says it should be sold for [↑](#footnote-ref-1)