Prepare attributes to upload in Django

Contents

[Introduction 2](#_Toc121296368)

[Attributes serialization 3](#_Toc121296369)

[File format for Attributes loading into Django 4](#_Toc121296370)

[Duplicates not allowed in attributes 5](#_Toc121296371)

[Example 5](#_Toc121296372)

[Example 5](#_Toc121296373)

[Example 5](#_Toc121296374)

[Attributes entities process rules 6](#_Toc121296375)

# Introduction

It is data processing intension task. Commonly it done in applications like Excel that supports advanced formulas or databases like Oracle that supports PL/SQL. You are free to use any tools you are familiar with to prepare attributes data files.

Django has 2 types of product attributes:

* **Attributes** that used for filtering products in Django.
* **Detail attributes** that used as product specifications in Django.

Basically, both Attributes and Detail attributes are product specifications. But Attributes

In order to prepare Attributes and Detail attributes for loading into Django you have to have following data exported from Django:

* .csv file with ID, CATALOG\_NUMBER, ATTRIBUTES fields.
* .csv file with ID, CATALOG\_NUMBER, DETAIL\_ATTRIBUTES fields.

# Attributes serialization

Attributes and Detail Attributes should be serialized into string.

Example of serialized string: “HP:2|Color:Green|Material:Stainless steel”

Where:

* HP, Color and Material are attribute names.
* 2, Green and Stainless steel are attribute values.

Attribute name should not contain “:” symbol. If such symbol is present then it should be replaced with any other appropriate symbol.

Attribute value should not contain “|” symbol. If such symbol is present then it should be replaced with any other appropriate symbol.

Attributes and Detail attributes have the same rules for serialization.

# File format for Attributes loading into Django

Attributes should be loaded into Django as a **.csv** file with comma (,) delimiter.

Attributes file should have only 2 fields:

* ID, contains product ID in Django.
* ATTRIBUTES, contains serialized attributes.

Attributes file should have only 2 fields:

* ID, contains product ID in Django
* DETAIL\_ATTRIBUTES, contains serialized attributes.

# Duplicates not allowed in attributes

Entities in Attributes shouldn’t duplicated in Detail attributes.

## Example

If we have Attributes like “HP:2|Color:Green|Material:Stainless steel”

If “HP:2” moved to Attributes. Then only “Color:Green|Material:Stainless steel” should be moved to Detail attributes.

Entities in newly scraped Attributes shouldn’t be duplicate already exists Attributes in Django.

## Example

If we have newly scraped Attributes like “HP:2|Voltage (V):380”. At same time, Django has already “HP:2” Attribute for the product. Then we should load into Django only “Voltage (V):380”.

Entities in newly scraped Detail attributes shouldn’t duplicate already exists Detail attributes in Django.

## Example

If we have newly scraped Detail attributes like “Color:Green|Material:Stainless steel”. At same time, Django has already “Color:Green” Detail attribute for the product. Then we should load into Django only “Material:Stainless steel”.

# Attributes entities process rules

**Attributes and Detail attributes have identical entities process rules.**

1. Attribute names

1.1 Attribute names have to start with uppercase letter. `Bore` is correct, `bore` is not correct.

1.2 Attribute names have not to contain measuring unit. `Bore` is correct, `Bore (in)` is not correct.

1.3 Every word in attribute names have to start with uppercase letter. `Bore Size` is correct, `Bore size` is not correct.

1.4 Attribute names have not to consist of only uppercase letters. `Bore` is correct, `BORE` is not correct.

2. Measurement units in attribute values

2.1 Use a space in attribute values between value and measure unit. `4 in` is correct, `4in` is not correct.

2.2 Use `in` over `"` for inch measure unit. `4 in` is correct, `4"` is not correct.

2.3 Use shorthands like `in`, `ft`, `mm` over full length words like `inches`, `foot`, `millimeters`. `4 in` is correct, `4 inches` is not correct.

2.4 Use imperial system convention over metric system convention. `4 in` is correct, `101.6 mm` is not correct.

2.5 Attribute values in metric system have to be converted to imperial system. `101.6 mm` convert to `4 in`.

2.6 Do not use both metric systems (imperial and metric) at the same time. `1-1/2 in` is correct, `1-1/2" (38.1 mm)` is not correct

2.7 Do not use special HTML symbols as measure unit. `15°` is correct, `15&deg;` is not correct.

2.8 Use only lowercase letters in measure unit names. `50 ft` is correct, `50 FT` is not correct.

3. Number format in attribute values

3.1 Use decimal point numbers over fractional numbers. `0.5 in` is correct, `1/2 in` is not correct.

3.2 Always use leading zero before decimal point. `0.5 in` is correct, `.5 in` is not correct.

4. List of values in attribute values

4.1 List of values have to be splitted with comma. `Motorized, Quill Input` is correct, `Motorized; Quill Input` is not correct.

4.1 Every item in list of values have to start with upper letter. `Motorized, Quill Input` is correct, `Motorized, quill input` is not correct.