Statistics on the tropes and films in TvTropes. Technical report

Rubén H. García-Ortega

23rd June 2018

1 Introduction

A trope can be defined as any recurring plot device, motif, archetype or cliché used by the authors of creative works to build their books, films, comics and videogames, because they help them to reach their goals, that might vary from interest, surprise, familiarity, entertainment, etc. Some tropes are broadly adopted, as the three acts structure, the hero's journey, the McGuffin and the Chekhov's Gun, usually academically studied and promoted, but there are thousand of not-so-widely used tropes as well, discovered by professionals and enthusiastic of the storytelling. The study of the tropes is organic, dynamic and extensive, for this reason our reference is a wiki called TvTropes, that collects thousand of descriptions and examples of tropes from 2014 until now. The semantic network of knowledge behind TvTropes is huge and complex; it relates hierarchies of tropes with their usage in creative works in the pop culture. In order to make it usable by the community, AAA extracted all their data to RDF tuples in the so-called dbtropes.

The goal os this report is to analyse the films and the tropes they use in dbtropes from a statistical point of view, so it can be used and cited in further researches. In the framework of the Open Data, the information found in the report is released under Creative Commons and the resources are publicly available in the repository.

In order to easy the reproducibility of the results, the report includes python code chunks, and the values are dynamically calculated using pweave, scipy, numpy and matplotlib. The resulting pdf file is compiled through latex.

2 Pre-processing the data

dbtropes can be found as a 5GB RDF tuples file in aaaa. The size of the file makes it hard to load in a RDF tool in order to extract information. For this reason, we processed the file with regular expressions in a python script so we have a JSON file with films and the name of the tropes they use.

The rdf file contains XXX lines of tuples, meanwhile the new JSON file has XXX lines.

In order to easy the data analysis we will use two data structures: a dictionary of films where the values are lists of the tropes they use and the reversed dictionary, where the keys are the tropes and the values are lists of films that they are used in. The class that handles the python code can be found in.

3 Descriptive analysis of the films regarding their tropes

Feature	Tropes in a film
Observations	5925
Minimum	1
Maximum	515
Mean	43.434
Median	29.0
Q1	16.0
Q2	29.0
Q3	52.0
Variance	2133.35
Skewness	3.332
Kurtosis	17.373

Table 1: Descriptive analysis

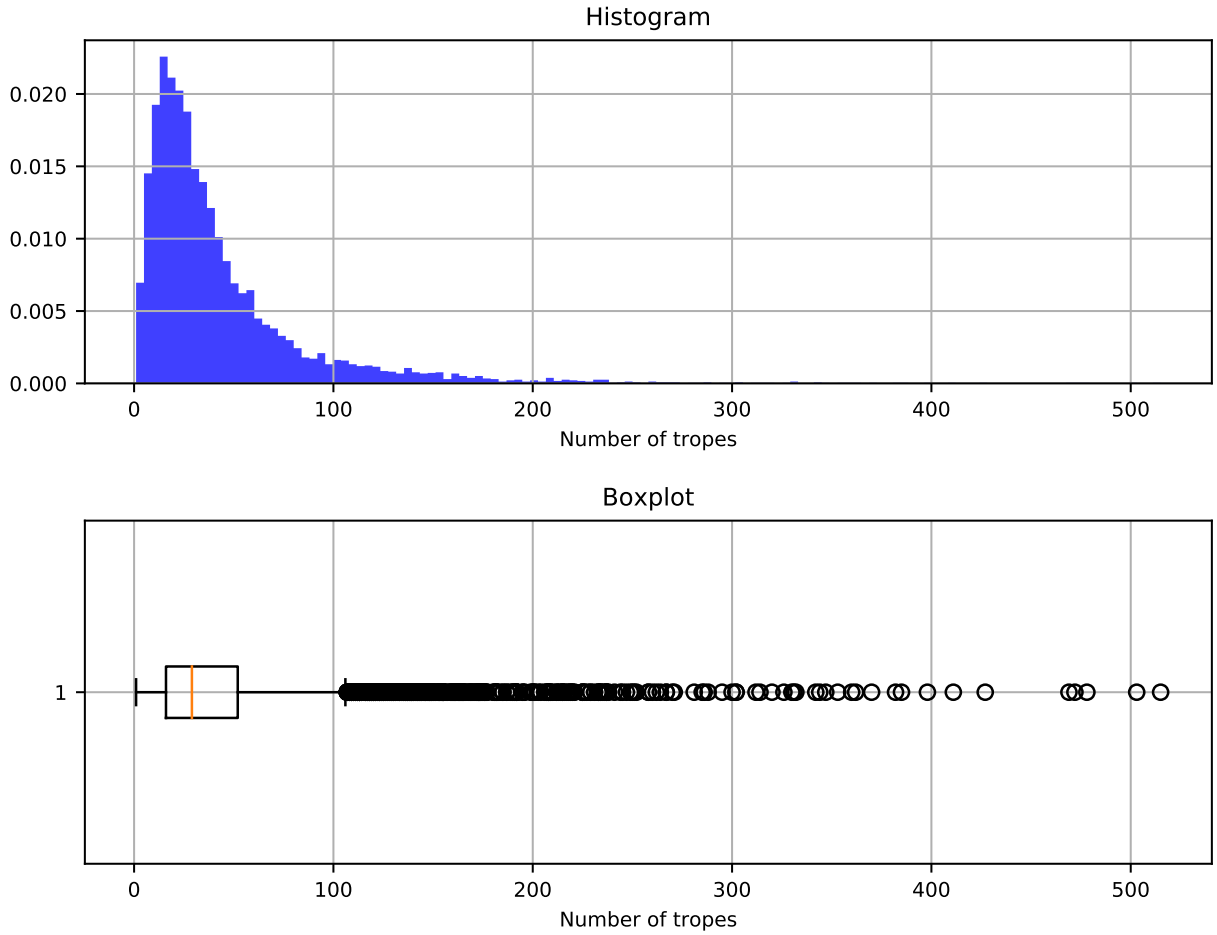

Figure 1: Distribution of films by number of tropes

Film	N. tropes
GuardiansOfTheGalaxy	515
TheDarkKnightRises	503
XMenDaysOfFuturePast	478
CaptainAmericaTheFirstAvenger	472
XMenFirstClass	469
Thor	427
SherlockHolmes	411
TheLordOfTheRings	398
PacificRim	385
CaptainAmericaTheWinterSoldier	382
WhoFramedRogerRabbit	370
TheDarkKnight	362
TronLegacy	360
StarTrek	353
StarTrekIntoDarkness	347
Skyfall	344
TheGodfather	342
JurassicWorld	332
Serenity	331
BackToTheFuture	330
Inception	326
IronMan3	320
AustinPowers	314
GalaxyQuest	312
ManOfSteel	302

Table 2: Top10 films by number of tropes that they use

4 Descriptive analysis of the tropes

Feature	Tropes by film
Observations	18270
Minimum	1
Maximum	1502
Mean	14.086
Median	5.0
Q1	2.0
Q2	5.0
Q3	12.0
Variance	1464.794
Skewness	11.758
Kurtosis	245.951

Table 3: Descriptive analysis

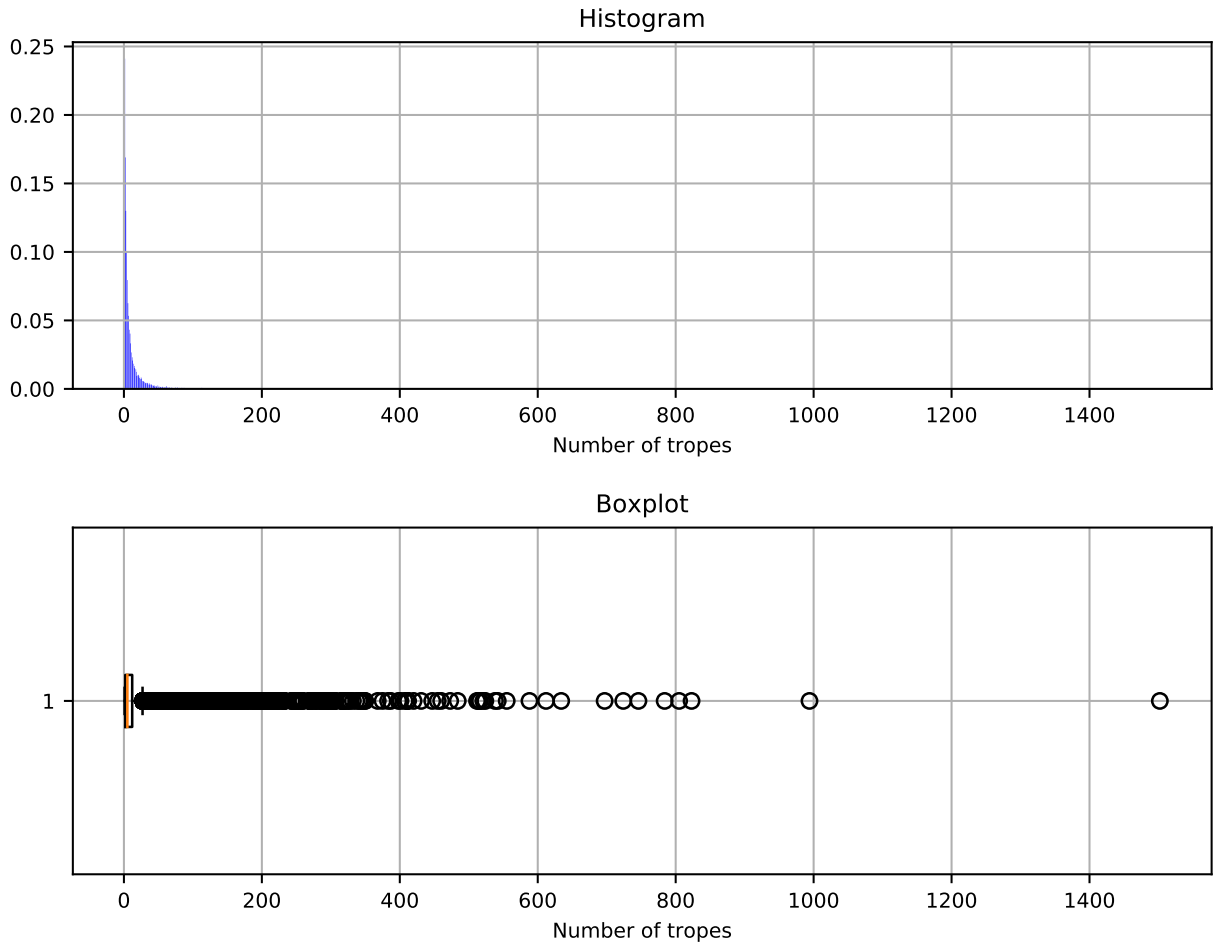

Figure 2: Distribution of tropes by number of films

Film	N. films
ShoutOut	1502
ChekhovsGun	994
OhCrap	823
DeadpanSnarker	805
Jerkass	784
Foreshadowing	746
LargeHam	724
BittersweetEnding	697
TitleDrop	634
BigBad	612
MeaningfulName	588
BerserkButton	555
TheCameo	542
WhatHappenedToTheMouse	538
RunningGag	524
TooDumbToLive	521
DownerEnding	516
FanService	516
KarmaHoudini	514
GroinAttack	512
BrickJoke	484
BookEnds	473
MoodWhiplash	460
KickTheDog	455
PrecisionFStrike	447

Table 4: Top10 tropes by number of films that they appear in