

Python for Practice – Homework 3

The goal of this homework is to implement several queries over the reviews/reviewers dataset. (The homework corresponds to exercises e06 – e08.)

Query #1

Print out 10 men and 10 women who wrote the longest reviews. Every person is to be displayed only once.

Sample output (note that your results or their actual representation may differ):

Top 10 men

id	name	maxReviewLength	gender	age
A1X4J08EJ1U5BR	Mike London "MAC"	12789	Male	41
A2582KMXLK2P06	B. E Jackson	10580	Male	60
A8EDTKSPOMRWK	hyperbolium	7893	Male	53
ANCOMAI0I7LVG	Andrew Ellington	6197	Male	64
AHN34TYB15NWF	! Metamorpho ;)	6121	Male	32
A6UW2KE05UOSA	Corey Turner "The...	6120	Male	22
A2AOZQ3WTVVOK	Lonnie E. Holder ...	5707	Male	58
A2I2A7N30CPNW4	T. A. Daniel "Alex"	5329	Male	56
A2WETFLLJM0D6R0	ewomack "ewomack"	5253	Male	60
A1QP5VCWEL1LHD	Humberto Mejia "S...	5230	Male	15

Top 10 women

id	name	maxReviewLength	gender	age
A39W3263A9HCMN	Renfield "Up the ...	8766	Female	16
A16QODENB3VUI1	Robert Moore	7399	Female	25
A388ZDFVK72VCA	Cale E. Reneau "M...	6732	Female	54
A11D1KHM7DVOQK	Person "qazwsx34651"	6211	Female	48
A3HU0B9XUEVHIM	Andre S. Grindle ...	5758	Female	48
A26NLSTT75FMJM	finulanu ""th...	5245	Female	60
A1CS70DT7ZPCJJ	Kasey G	5185	Female	33
A1D2C0WDCSHUWZ	E. A Solinas "ea_...	4931	Female	65
A9Q28YTLYRE07	mistermaxxx08 "mi...	4929	Female	39
ACVFTS9X2MKJ8	Rykre "The Rogue ...	4837	Female	43

Query #2

Compute the number of reviews per reviewer. Similar to e05, make a statistics with breakdown by gender and age.

Sample output (note that your results or their actual representation may differ):

gender	ageRange	avg(reviewsCount)	min(reviewsCount)	max(reviewsCount)	stddev_samp(reviewsCount)
null	null..null	27.241935483870968	10	334	35.64753226008293
null	10..19	19.944444444444443	10	66	13.739047925838847
null	20..29	25.6	10	147	26.209500778047698
null	30..39	34.5531914893617	10	334	54.988855989838804
null	40..49	27.81578947368421	10	175	34.260413893744726
null	50..59	20.777777777777778	10	62	13.63912171019971
null	60..69	26.846153846153847	10	132	28.214453470081335
Female	null..null	28.214285714285715	10	334	43.768303277289675
Female	10..19	22.6	10	66	17.1671520966965
Female	20..29	30.928571428571427	10	147	35.95823279159455
Female	30..39	34.470588235294116	10	334	77.52912166329729
Female	40..49	28.352941176470587	10	175	40.40566355176986
Female	50..59	20.857142857142858	10	49	13.999215048638334
Female	60..69	29.25	11	132	34.16836655046994
Male	null..null	26.441176470588236	10	178	27.42102016915244
Male	10..19	16.625	10	29	7.539183358579741
Male	20..29	20.9375	10	49	12.561681150758975

Male	30.39	34.6	10	178	38.47130592156114
Male	40.49	27.38095238095238	10	135	29.40999182331779
Male	50.59	20.692307692307693	10	62	13.810531098794469

Query #3

Select and manually print the 15 newest reviews on items that are ranked on average above 4.0. The reviews displayed should be only those that are ranked helpful above average over all reviews (this is the first element in the "helpful" array). To select the first element of the array, use the array indexing syntax - i.e. `reviewDF.helpful[0]`

Sample output (note that your results or their actual representation may differ):

```

Top 15:
12 20, 2013 AC0Z6U90WFRT0 B00BB22GQM Daniel Maltzman 3.0 4 The award for most overhy-
ped/overrated album of the year goes to...
11 5, 2013 AX3WJZLFQQRCT B00F1CRRIU R. M. Ettinger "rme1963" 3.0 4 Excess is not Enough
11 4, 2013 A3HU0B9XUUEVHIM B00F1CRRIU Andre S. Grindle "Andre' Grindle" 5.0 4 Potentially
Innovative Music From A Potentially Innovative Band
10 30, 2013 AXHWQ834G3IUV B00F1CRRIU Paul D. Sandor 3.0 12 Shiny flawed diamond...
10 30, 2013 ALND45MWBGUQS B00F1CRRIU Lunar Boulevard 3.0 18 I Know We Can Find a Way...
10 22, 2013 A2ULSA0IRUE09E B000002BE9 None 5.0 4 This Only Happens Once
05 20, 2013 A14GK0E64J0WAS B00BB22GQM Paul Allaer 3.0 4 Goodbye fun songs, hello "matu-
rity"
05 14, 2013 AE31M52VLKOG6 B00BB22GQM Robert G Yokoyama 5.0 4 Modern Vampires of the City
is an entertaining album.
05 14, 2013 A2I2A7N30CPNW4 B00BB22GQM T. A. Daniel "Alex" 5.0 5 Early Contender for One of
2013's Best
04 6, 2013 A1XP7P3X9MRE5H B000003BCX Morten Vindberg 4.0 4 Late-night Blues
03 8, 2013 A1429V2M057X86 B00AA0TZW JUDAS TREES 5.0 4 no complaints
03 5, 2013 A3HU0B9XUUEVHIM B00AA0TZW Andre S. Grindle "Andre' Grindle" 5.0 5 May Be The
Umpteenth Unreleased Hendrix Album,But Still Worthy Of Praise
03 5, 2013 A4TUG0Q605QFO B00AA0TZW Dave &#34;Fever Tree&#34; Sigmon 3.0 10 Experience
Hendrix's Barrel Scrappings
03 5, 2013 A2P0PLCR7JWwNR B00AA0TZW James E. Lane "jaminjim" 2.0 15 Excuse me while I
skip this one...
12 10, 2012 A5SC908TJZJ6DK B00900PBTC Gregory William Locke 3.0 6 Good Enough, Sure, But
Hardly the Year's Best Hip-Hop Record (That Belongs to Joey Bada$$)

```