# Quran Statistics

RDS-Q No. #4 | **Verse Data (6236 Verses)**

# Contents

# Overview

The Holy Book Quran consists of **114** Chapters (Surah) and totally **6236** Verses (Ayah). The Quran was orally revealed over a period of 23 years and written for the first time at the time of Khulafaur Rasyidin. In the most printed editions, the Quran consists of 604 pages, which furtherly organized into parts, notably a so-called Juz. There is 30 Juz in total.

For the results presented in this document, the Quran text is based on the Uthmani version published by Tanzil project (http://tanzil.net/). Based on this text, the Quran is composed of **77430** words and **325666** letters (note: un-numbered Basmallah at the beginning of 112 chapters is not counted). For comparison, based on data published by corpus.quran.com (Kais Dukes, University of Leeds) the figures are **77429** words and **623638** (join letters and diacritics/harakah). The only difference is at QS 37:130 for the arabic words إِلْ يَاسِينَ (trans: Prophet Elijah / Ilyas a.s.). Quran corpus counts it as one word, while it in Tanzil's version of the Uthmani text is written as two words.

This document mainly presents some numbers and figures with minimal narration since it is meant to be a quick reference for supporting further research in some aspects of The Noble Quran. Data used in RDS-Q #4 is at **verse level** (length of **6236**) i.e. data related to words and letters are pre-processed for the each verse, while data for juz and page is aggregated correspondingly.
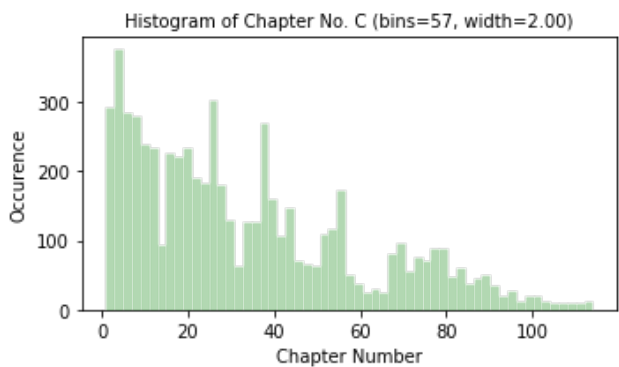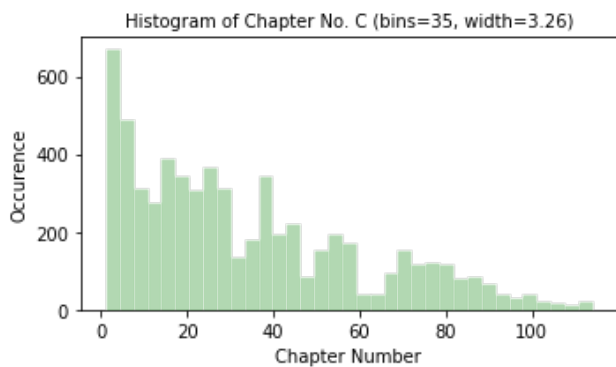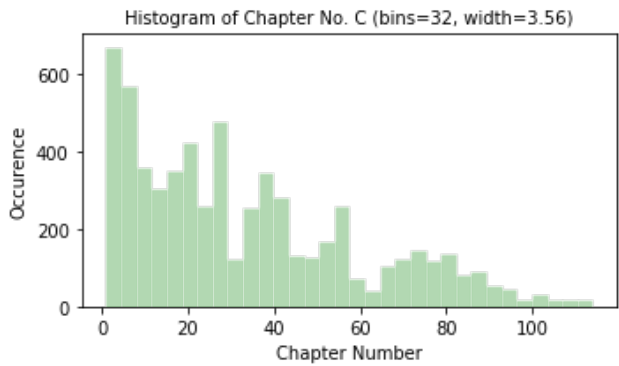
# All Verses

## Chapters & Verses

The 6236-length verse data is formed orderly by sequential chapter number (C) and sequential verse number (V). The Figure below clearly shows this arragement. The change in C is marked by resetting V to the base value of 1, forming a sawtooth-like graph. Length of the chapter is the distance beetween two points at this base value in the x-axis.
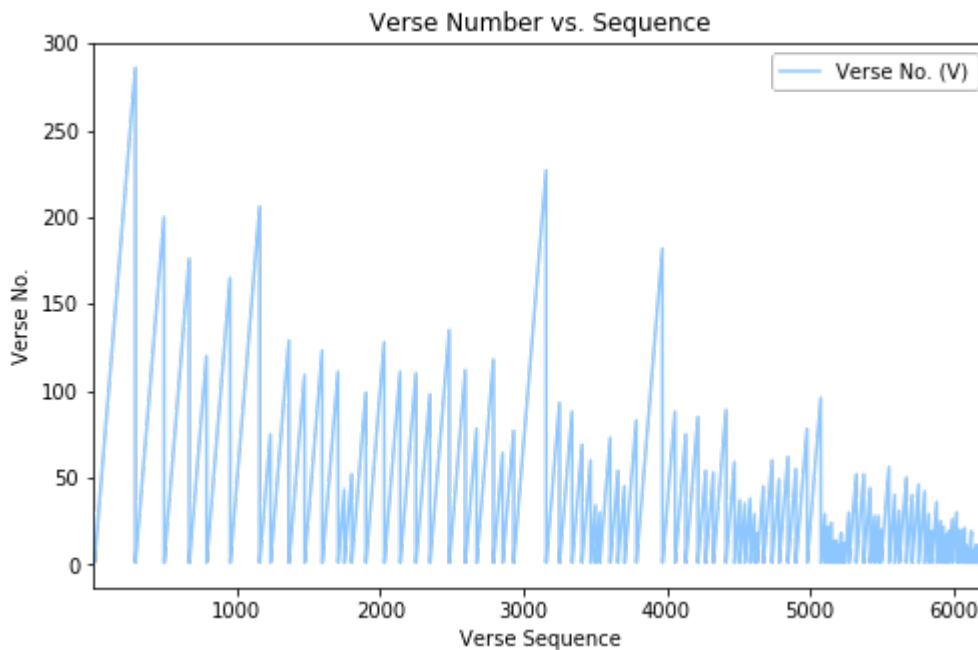


The following Figures give some histogram plots for C, V and the value of C+V. For each Figure, the number of bins are varied to see the behaviour of the parameters in different number of groups. While all parameters have shown some degree of possible patterns, the most interesting ones can be seen in the histograms of V and C+V.

Histogram of Verse No. V (bins=29, width=9.86)

Histogram of Verse No. V (bins=57, width=5.02)

Histogram of Verse No. V (bins=86, width=3.33)

Histogram of Verse No. V (bins=114, width=2.51)

Histogram of Chapter No. C (bins=29, width=3.93)

Histogram of Chapter No. C (bins=32, width=3.56)

Histogram of Chapter No. C (bins=35, width=3.26)

Histogram of Chapter No. C (bins=57, width=2.00)
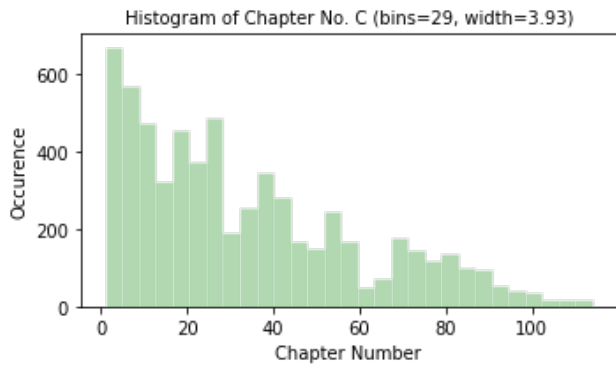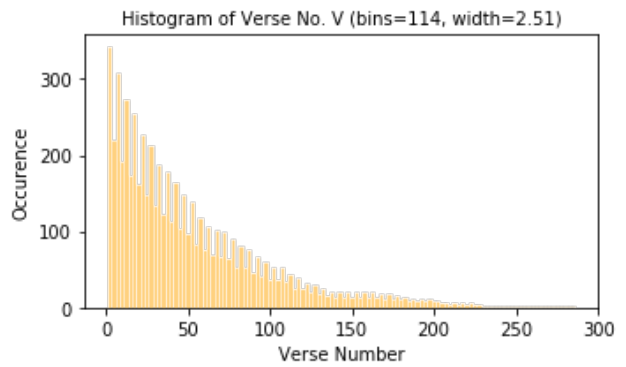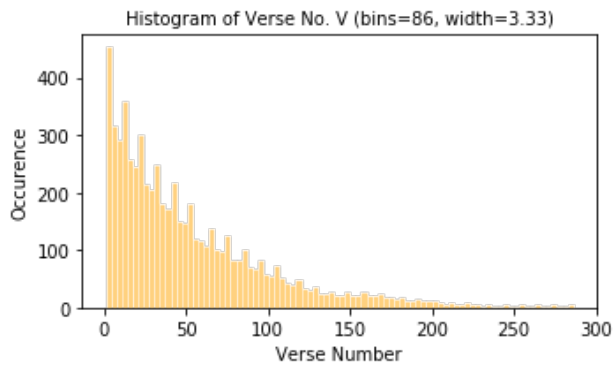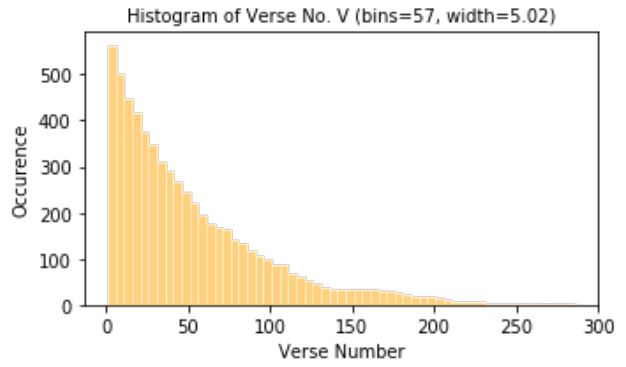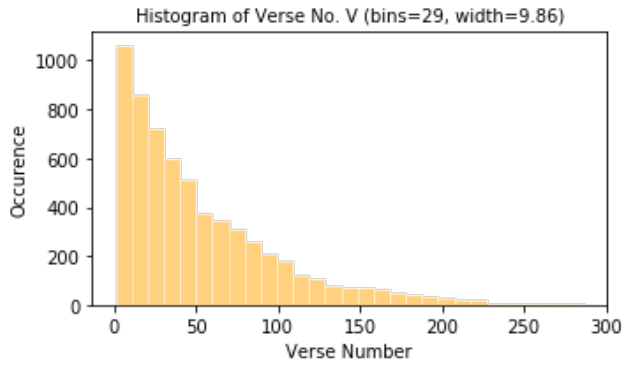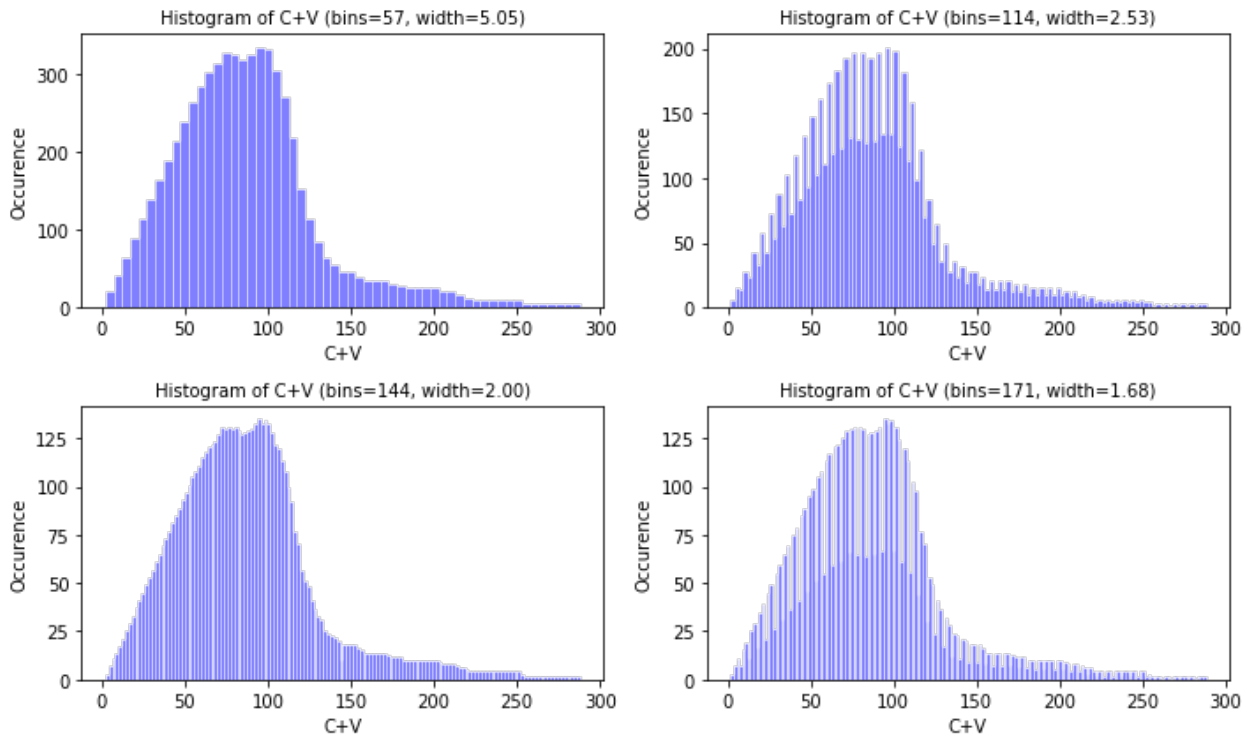
The last Figure in this section depicts standard density functions for each parameters.
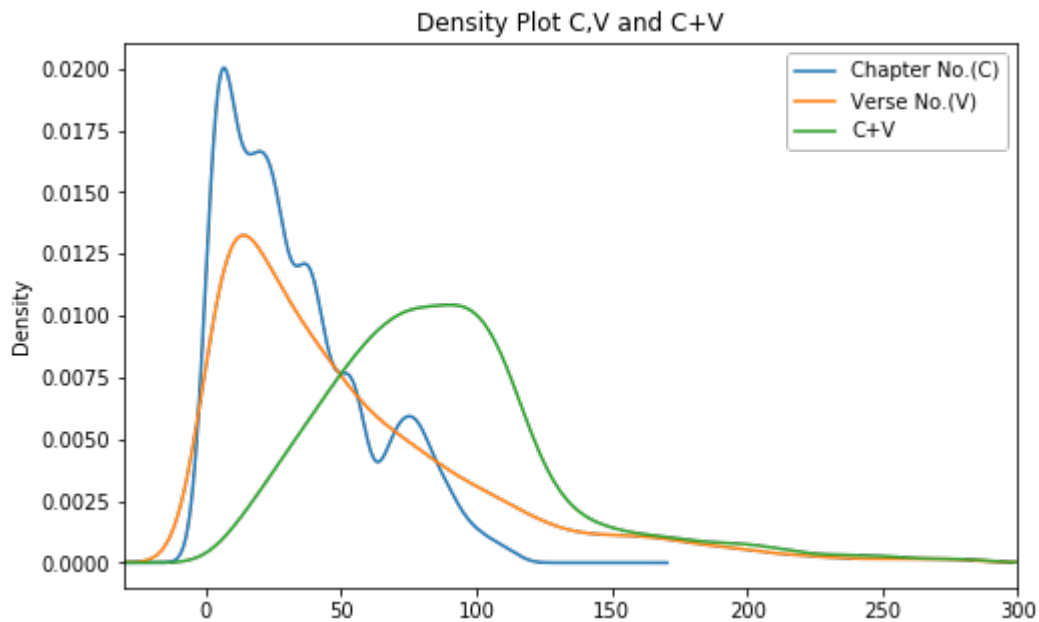
The following Table gives some statistical values for C,V and C+V parameters.

|  | MAX | 75% | 50% | 25% | MIN | MEAN | STD | COUNT | SUM |
|---|---|---|---|---|---|---|---|---|---|
| Chapter No. (C) | 114 | 51.0 | 26.0 | 11.0 | 1 | 33.52 | 26.46 | 6236 | 209029 |
| Verse No. (V) | 286 | 75.0 | 38.0 | 16.0 | 1 | 53.51 | 50.46 | 6236 | 333667 |
| C+V | 288 | 107.0 | 83.0 | 58.0 | 2 | 87.03 | 43.98 | 6236 | 542696 |

# Words & Letters

The following five Figures focus on the number of words (W) and letters (L).



Number of Words for each Verse



Histogram of Num. of Words W (bins=25, width=5.12)



Histogram of Num. of Words W (bins=50, width=2.56)



Histogram of Num. of Words W (bins=75, width=1.71)



Histogram of Num. of Words W (bins=100, width=1.28)
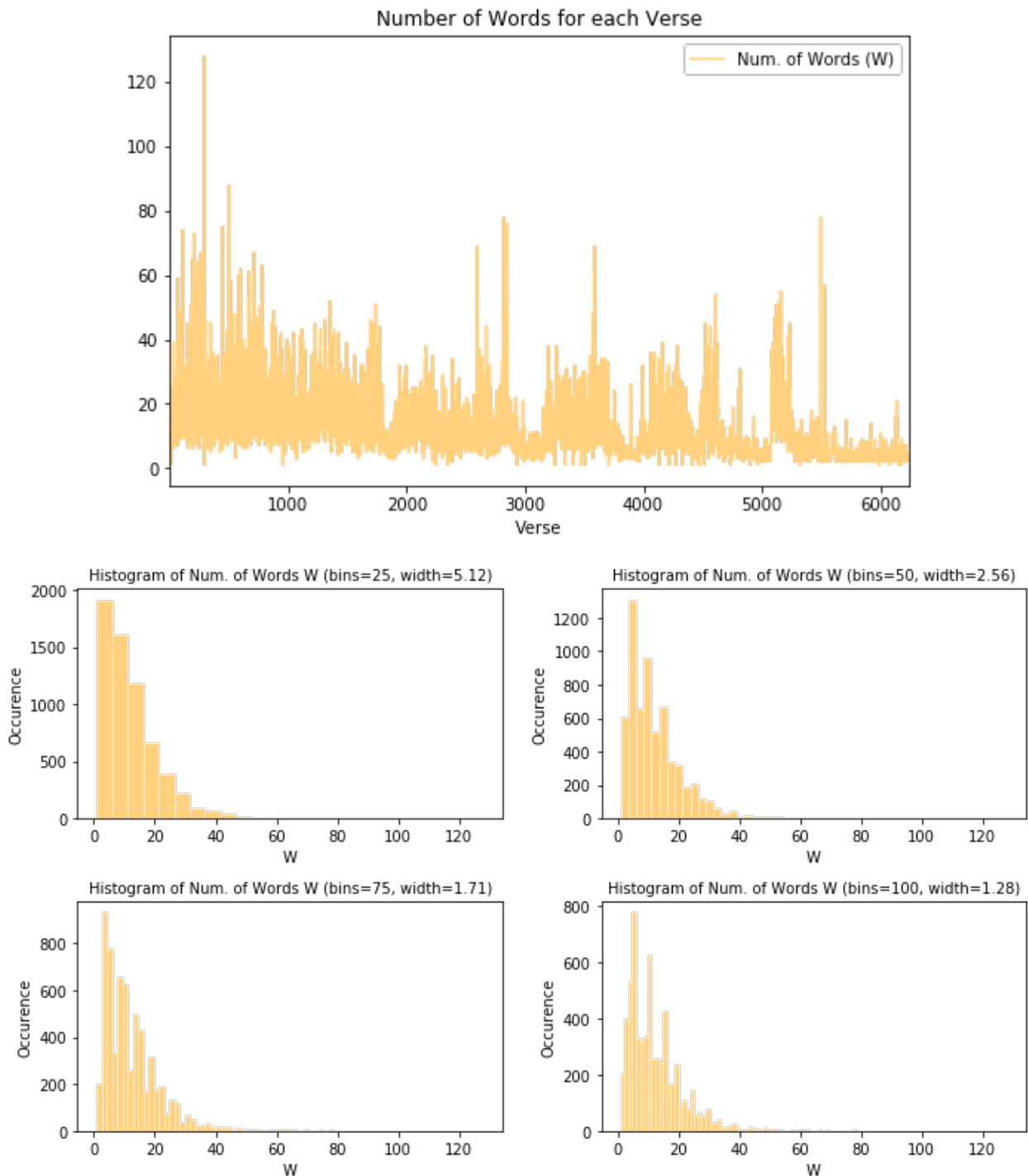
The first Figure depicts linear plot of W over verse sequence. The second Figure gives some higtogram of W for different number of bins.



Number of Letters for each Verse



Histogram of Num. of Letters L (bins=25, width=21.88)



Histogram of Num. of Letters L (bins=50, width=10.94)



Histogram of Num. of Letters L (bins=100, width=5.47)



Histogram of Num. of Letters L (bins=125, width=4.38)
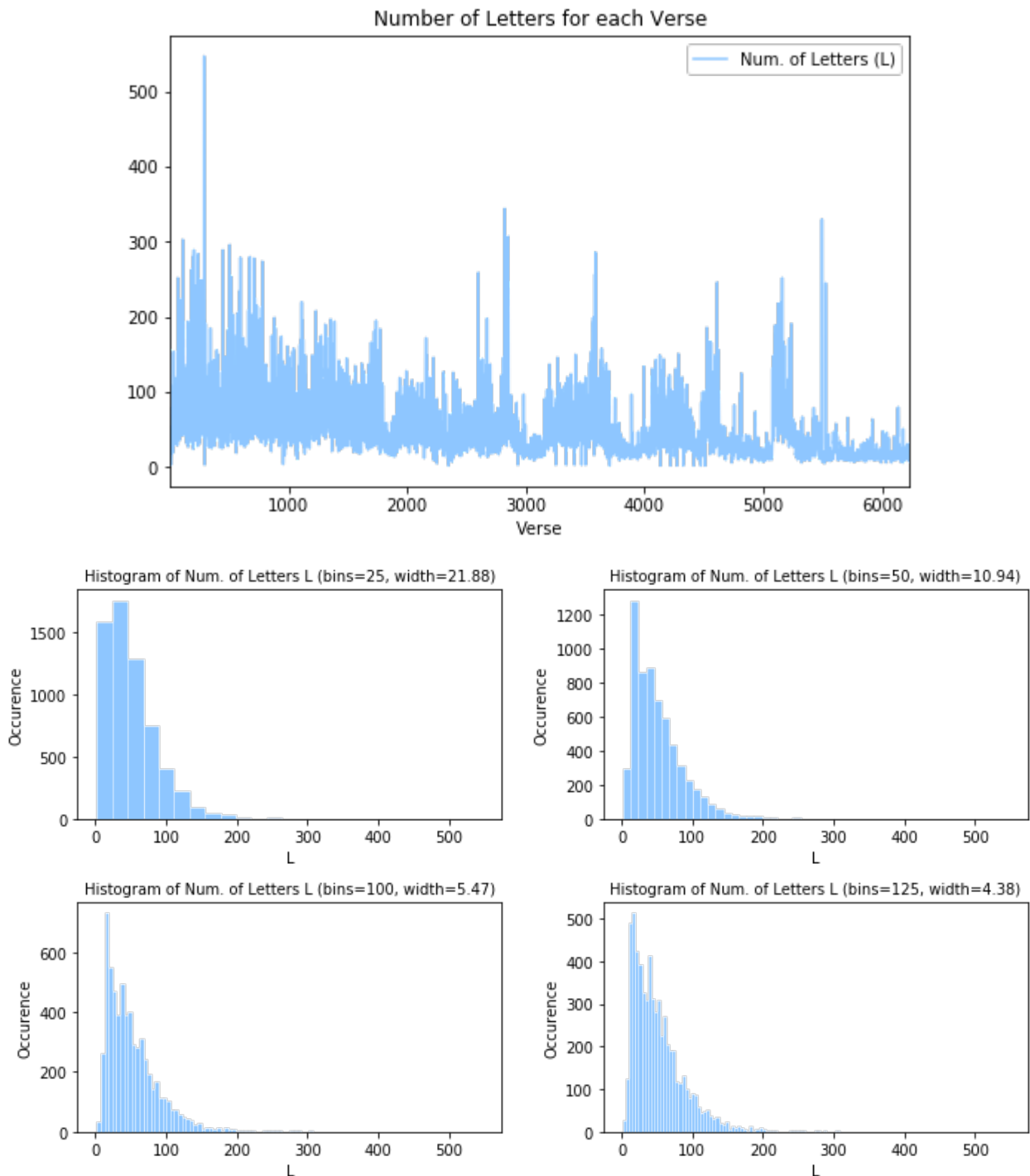
The third Figure depicts linear plot of L over verse sequence. The fourth Figure gives some higtogram of L for different number of bins. The last Figure below shows density plot of both W and L.

### Density Plot W and L
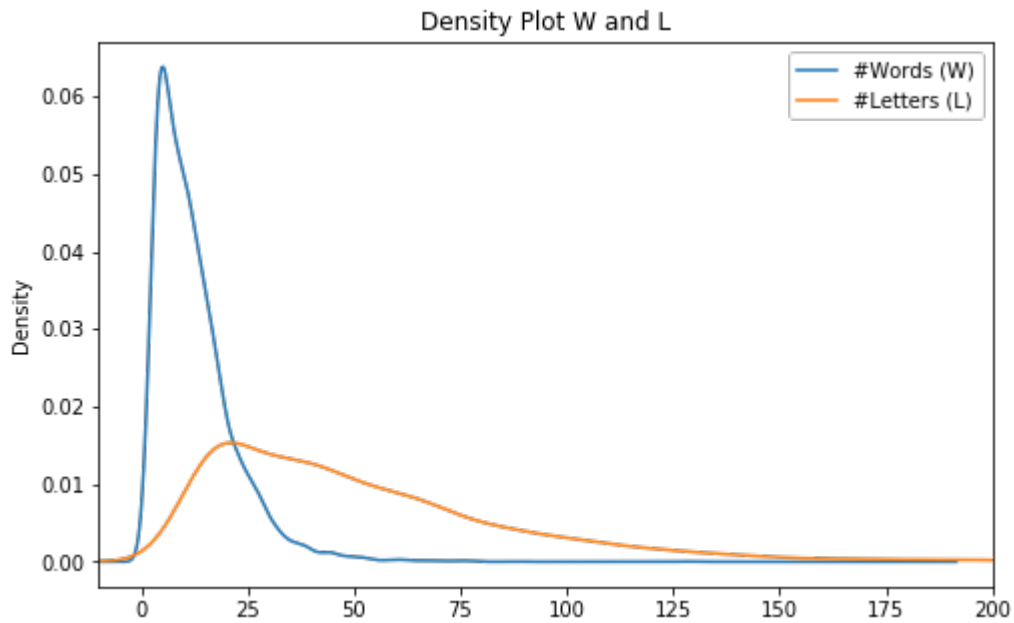


The following Table gives some statistical values for W and L.

|  | MAX | 75% | 50% | 25% | MIN | MEAN | STD | COUNT | SUM |
|---|---|---|---|---|---|---|---|---|---|
| Num. of Words | 128 | 16.0 | 10.0 | 6.0 | 1 | 12.42 | 9.42 | 6236 | 77430 |
| Num. of Letters | 547 | 68.0 | 43.0 | 23.0 | 2 | 52.22 | 39.25 | 6236 | 325666 |

# Data Split

# [1-3118] & [3119-6236]

For some investigations we might want to look at a segment of the data. The following Figure shows density functions of C,V and C+V for the segment [1-3118] and [3119-6236].
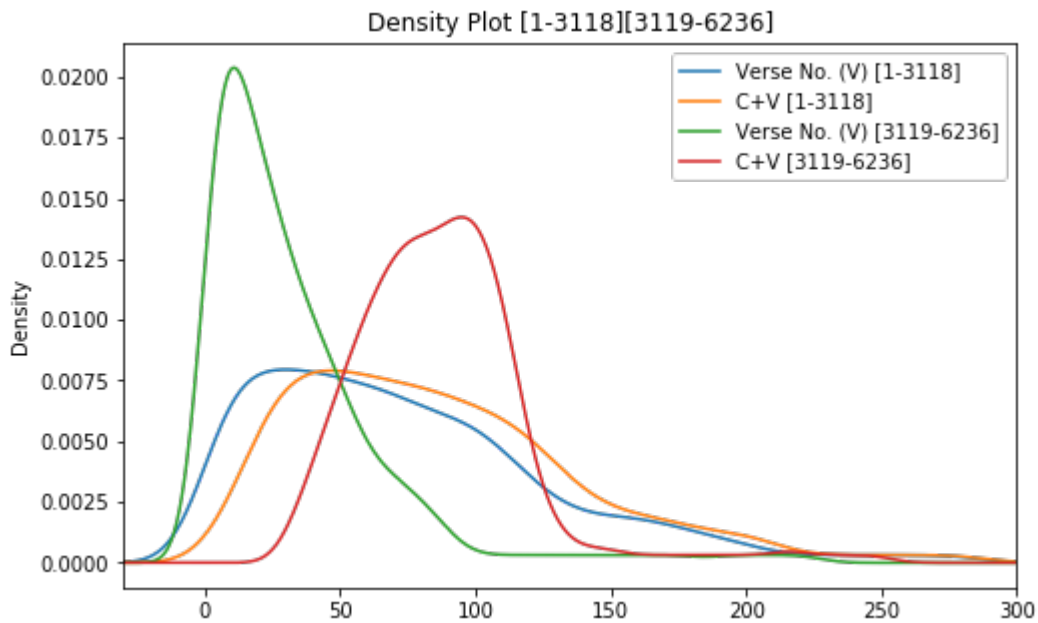


The following Table gives some statistical values for C,V and C+V for segmen [1-3118].

|  | MAX | 75% | 50% | 25% | MIN | MEAN | STD | COUNT | SUM |
|---|---|---|---|---|---|---|---|---|---|
| Chapter No. (C) | 26 | 19.0 | 11.0 | 5.0 | 1 | 12.44 | 7.8 | 3118 | 38779 |
| Verse No. (V) | 286 | 103.0 | 64.0 | 31.0 | 1 | 74.05 | 55.08 | 3118 | 230879 |
| C+V | 288 | 116.0 | 78.0 | 45.0 | 2 | 86.48 | 53.19 | 3118 | 269658 |

The following Table gives some statistical values for C,V and C+V for segmen [3119-6236].

|  | MAX | 75% | 50% | 25% | MIN | MEAN | STD | COUNT | SUM |
|---|---|---|---|---|---|---|---|---|---|
| Chapter No. (C) | 114 | 72.0 | 51.0 | 37.0 | 26 | 54.6 | 21.23 | 3118 | 170250 |
| Verse No. (V) | 227 | 44.0 | 23.0 | 10.0 | 1 | 32.97 | 34.88 | 3118 | 102788 |
| C+V | 253 | 103.0 | 85.0 | 67.0 | 28 | 87.57 | 32.26 | 3118 | 273038 |



Density Plot [1-3118][3119-6236]
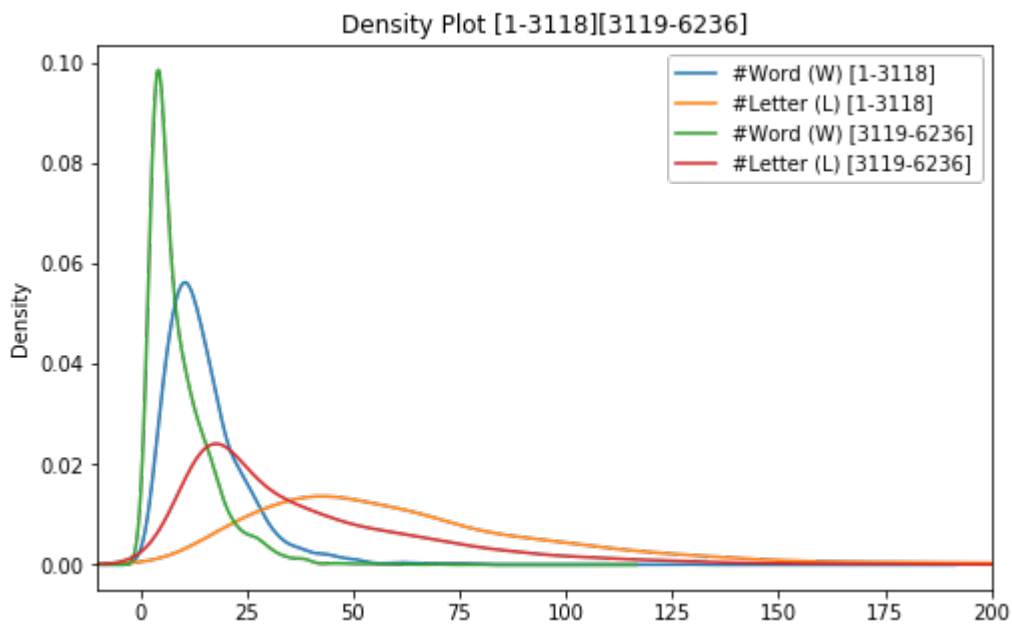
The following Table gives some statistical values for W and L for segmen [1-3118].

|  | MAX | 75% | 50% | 25% | MIN | MEAN | STD | COUNT | SUM |
|---|---|---|---|---|---|---|---|---|---|
| Num. of Words | 128 | 19.0 | 13.0 | 9.0 | 1 | 15.46 | 10.02 | 3118 | 48210 |
| Num. of Letters | 547 | 82.0 | 55.0 | 37.0 | 2 | 65.0 | 41.79 | 3118 | 202681 |

The following Table gives some statistical values for W and L for segmen [3119-6236].

| | MAX | 75% | 50% | 25% | MIN | MEAN | STD | COUNT | SUM |
|---|---|---|---|---|---|---|---|---|---|
| Num. of Words | 78 | 13.0 | 7.0 | 4.0 | 1 | 9.37 | 7.64 | 3118 | 29220 |
| Num. of Letters | 330 | 53.0 | 29.0 | 17.0 | 2 | 39.44 | 31.75 | 3118 | 122985 |

# Odd/Even Split

# Chapters & Verses

For symmetry investigations we might want to look at a segment of the data which is based on certain criteria, in this case: odd/even C+V criteria as reported in RDS-Q #1 and RDS-Q #2. The following Figure shows density functions of C,V and C+V for the odd and even segment.
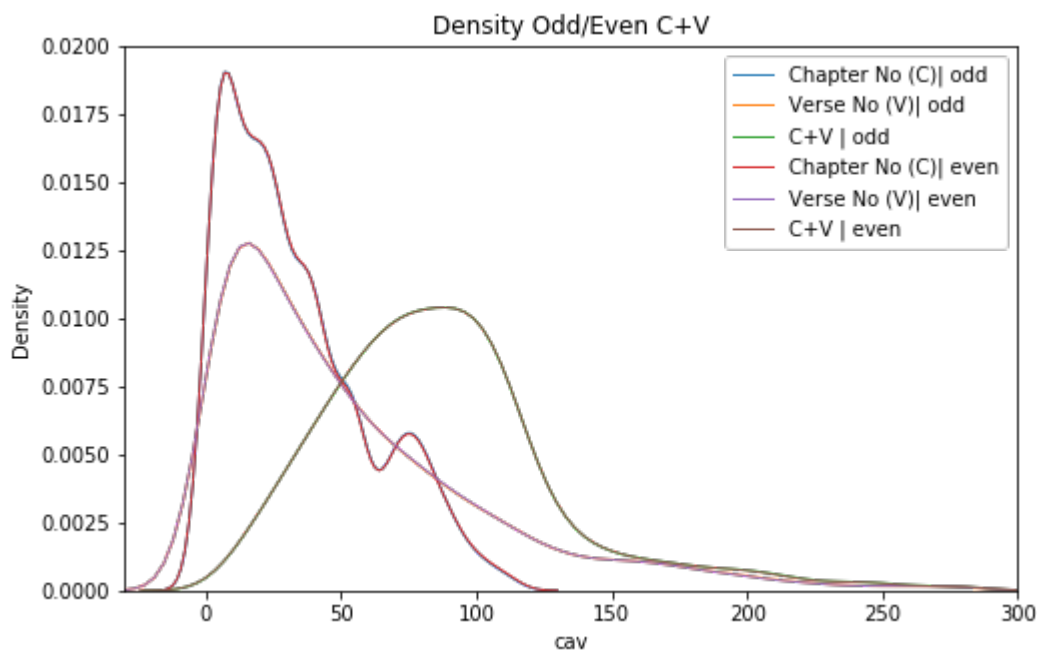


As shown above, the odd and even parameters share (almost) the same density curve. This is due to the nature of sequential number of both chapter number C and verse number V (within the same chapter). From the shape perspective, all curves are similar to those without data split except that the amplitude is lower.

The following Table gives some statistical values for C,V and C+V for **odd** segmen.

| | MAX | 75% | 50% | 25% | MIN | MEAN | STD | COUNT | SUM |
|---|---|---|---|---|---|---|---|---|---|
| Chapter No. (C) | 114 | 51.0 | 26.0 | 11.0 | 1 | 33.52 | 26.43 | 3118 | 104516 |
| Verse No. (V) | 285 | 75.0 | 38.0 | 16.0 | 1 | 53.54 | 50.48 | 3118 | 166926 |
| C+V | 287 | 107.0 | 83.0 | 57.0 | 3 | 87.06 | 44.0 | 3118 | 271442 |

The following Table gives some statistical values for C,V and C+V for **even** segmen.

| | MAX | 75% | 50% | 25% | MIN | MEAN | STD | COUNT | SUM |
|---|---|---|---|---|---|---|---|---|---|
| Chapter No. (C) | 114 | 50.75 | 26.0 | 11.0 | 1 | 33.52 | 26.49 | 3118 | 104513 |
| Verse No. (V) | 286 | 75.0 | 38.0 | 16.0 | 1 | 53.48 | 50.45 | 3118 | 166741 |
| C+V | 288 | 106.0 | 82.0 | 58.0 | 2 | 87.0 | 43.97 | 3118 | 271254 |

# Words & Letters

The following Figure shows density functions of W and L for the odd and even C+V segment.
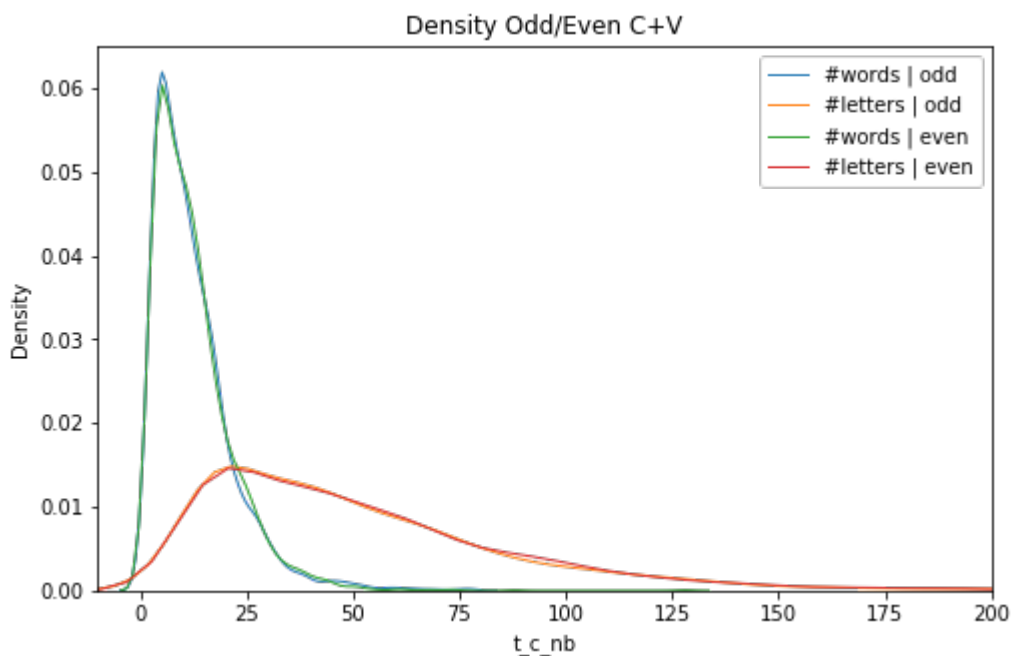


It is quite interesting that the density curves of both W and L for odd/even segment are almost identical. Differ to sequential parameter C and V previously mentioned, W and L are entirely derived from the verse text. The values in the table below are also interesting and have to some extent justified these curves.

The number of words for the odd/even group is **38716** / **38714**. The number of letters for the odd/even group is **162821** / **162845**. Both sum of W and L are almost halved. Note that if we are using data from corpus.quran.com, the number of words W for the odd group is **38715**, while that of the even group is **38714**.

The following Table gives some statistical values for W and L for **odd** segmen.

|  | MAX | 75% | 50% | 25% | MIN | MEAN | STD | COUNT | SUM |
|---|---|---|---|---|---|---|---|---|---|
| Num. of Words | 78 | 16.0 | 10.0 | 5.0 | 1 | 12.42 | 9.54 | 3118 | 38716 |
| Num. of Letters | 344 | 68.75 | 42.0 | 23.0 | 2 | 52.22 | 39.92 | 3118 | 162821 |

The following Table gives some statistical values for W and L for **even** segmen.

|  | MAX | 75% | 50% | 25% | MIN | MEAN | STD | COUNT | SUM |
|---|---|---|---|---|---|---|---|---|---|
| Num. of Words | 128 | 16.0 | 10.0 | 6.0 | 1 | 12.42 | 9.3 | 3118 | 38714 |
| Num. of Letters | 547 | 68.0 | 43.0 | 24.0 | 2 | 52.23 | 38.58 | 3118 | 162845 |

# Pages

## Chapters & Verses

As briefly mentioned in the overview, verse level data can be aggregated according to pages or juz. The Figure below depicts the number of chapters and verses for each page. Sure, the most pages contain only verses of a single chapter. The sum of the number of chapters includes repetitions.



Number of C,V for each Page
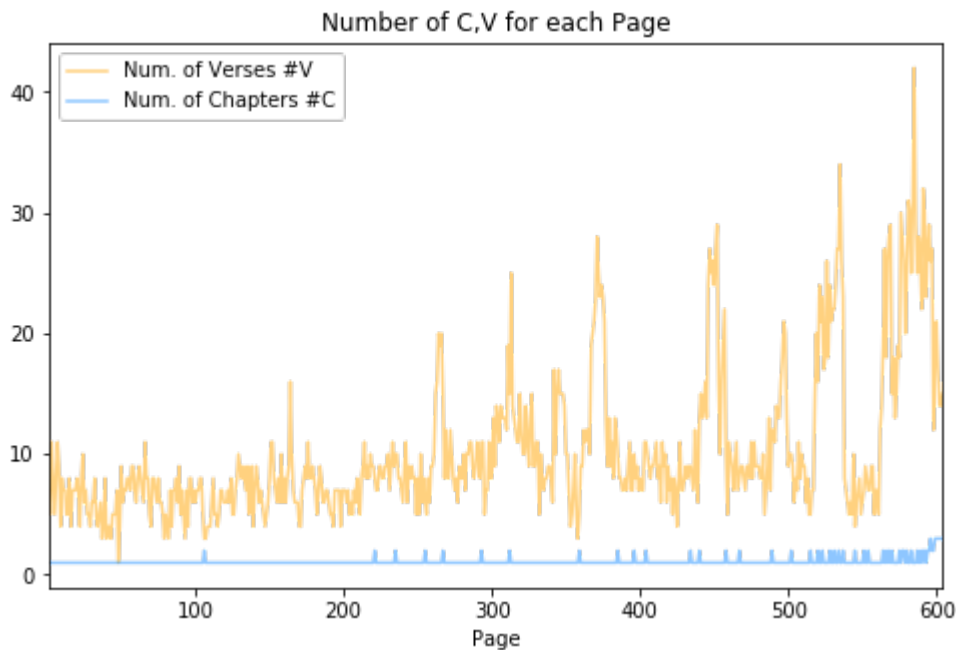
The following Table gives some statistical values for C and V for pages data.

|  | MAX | 75% | 50% | 25% | MIN | MEAN | STD | COUNT | SUM |
|---|---|---|---|---|---|---|---|---|---|
| Num. of Chapters | 3 | 1.0 | 1.0 | 1.0 | 1 | 1.1 | 0.33 | 604 | 662 |
| Num. of Verses | 42 | 11.0 | 8.0 | 7.0 | 1 | 10.32 | 6.18 | 604 | 6236 |

## ■ **Words & Letters**

Data can be further view to the level of words and letters. The following Figure gives the number of words and letters for each page.



Number of Words, Letters for each Page

The following Table gives some statistical values for W and L for pages data.

| | MAX | 75% | 50% | 25% | MIN | MEAN | STD | COUNT | SUM |
|---|---|---|---|---|---|---|---|---|---|
| Num. of Words | 161 | 137.0 | 129.0 | 121.0 | 29 | 128.2 | 14.92 | 604 | 77430 |
| Num. of Letters | 693 | 573.0 | 543.0 | 514.0 | 139 | 539.18 | 58.83 | 604 | 325666 |

# Juz

## Pages, Chapters & Verses

The following two Figures illustrate the number of pages (P), chapters (C) and verses (V) for each Juz.



Number of Pages, Chapters / Juz



Number of Verses / Juz
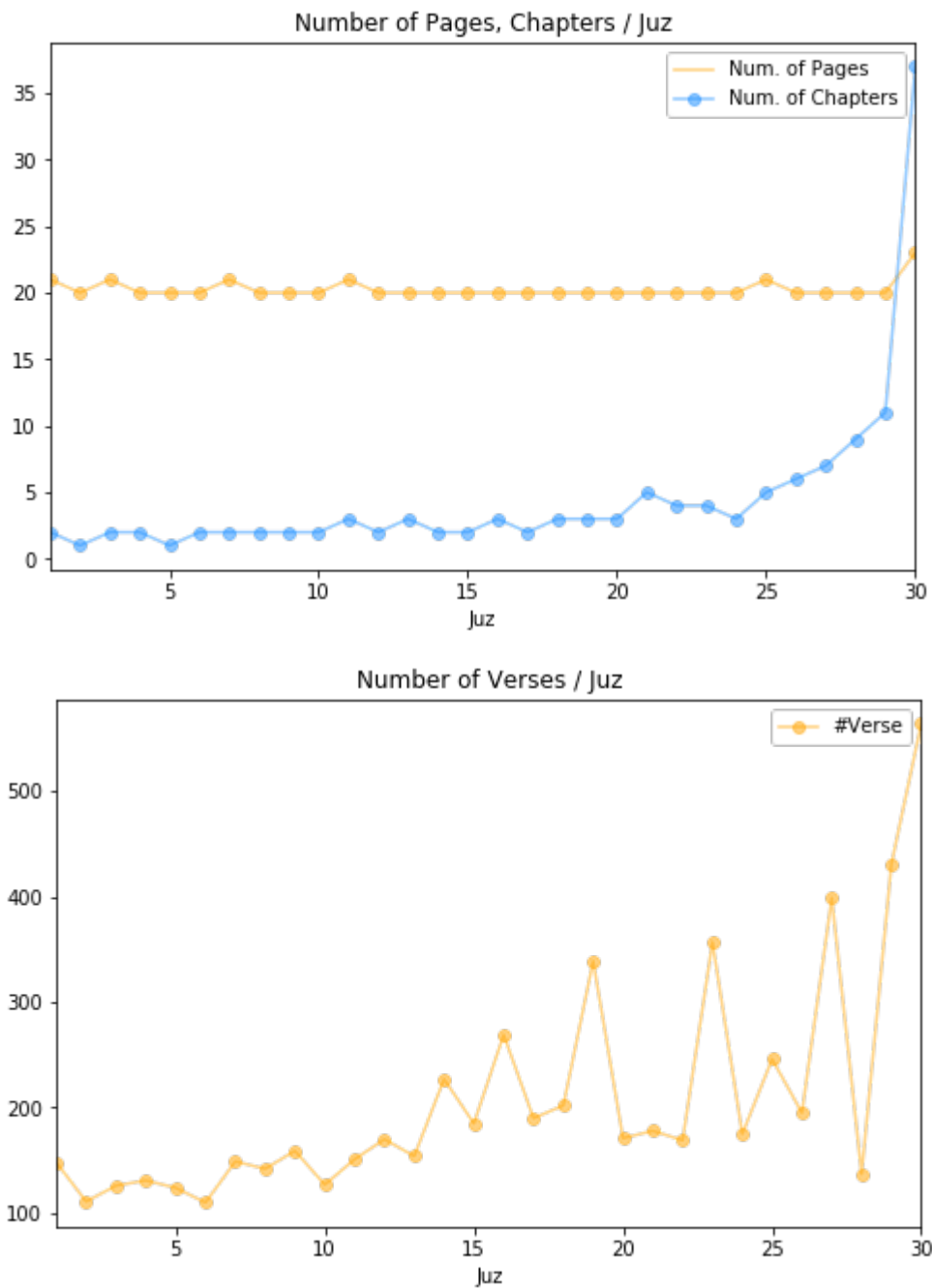
The plot of P and C are given in the first Figure, while the number of V is in the second one. As we might already be realized, the most of Juz consists of 20 pages.

The following Table gives some statistical values for P, C and V for Juz data.

| | MAX | 75% | 50% | 25% | MIN | MEAN | STD | COUNT | SUM |
|---|---|---|---|---|---|---|---|---|---|
| Num. of Pages | 23 | 20.0 | 20.0 | 20.0 | 20 | 20.27 | 0.64 | 30 | 608 |
| Num. of Chapters | 37 | 4.0 | 3.0 | 2.0 | 1 | 4.5 | 6.55 | 30 | 135 |
| Num. of Verses | 564 | 220.75 | 170.5 | 143.5 | 110 | 207.87 | 107.57 | 30 | 6236 |

Note that a chapter can span Juz boundaries and a page can contain end/start of Juz simultaneously (i.e. page: 62, 121, 201, 502).

## Words & Letters

The following Figure gives the number of words and letters for each Juz.
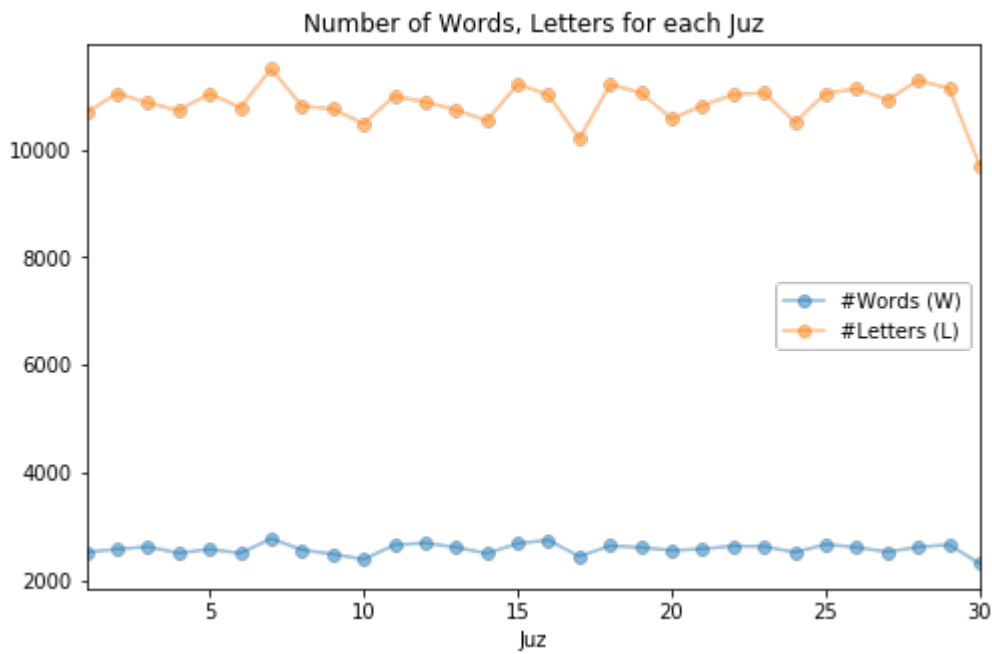


The following Table gives some statistical values for W and L for Juz data.

| | MAX | 75% | 50% | 25% | MIN | MEAN | STD | COUNT | SUM |
|---|---|---|---|---|---|---|---|---|---|
| Num. of Words | 2774 | 2640.25 | 2596.5 | 2520.5 | 2308 | 2581.0 | 100.27 | 30 | 77430 |
| Num. of Letters | 11497 | 11050.0 | 10900.0 | 10727.75 | 9704 | 10855.53 | 349.25 | 30 | 325666 |

# Resources

# ◼ **Resources**

| | URL |
|---|---|
| Data | https://github.com/eueung/rds-q/tree/master/data |
| PDF | https://github.com/eueung/rds-q/tree/master/PDF |
| Project (All) | https://github.com/eueung/rds-q/ |
| Web | https://quran.telematika.org/00004/quran-statistics-6236.html |
| Web (All) | https://quran.telematika.org/ |

**Sample Data**

| VE_NO_G | CH_NO | VE_NO | PAGE | JUZ | T_W_NB | T_C_NB | CAV | CAVOE |
|---|---|---|---|---|---|---|---|---|
| 2683 | 23 | 10 | 342 | 18 | 3 | 14 | 33 | odd |
| 3296 | 28 | 44 | 391 | 20 | 13 | 51 | 72 | even |
| 125 | 2 | 118 | 18 | 1 | 24 | 103 | 120 | even |
| 1 | 1 | 1 | 1 | 1 | 4 | 19 | 2 | even |
| 1895 | 15 | 93 | 267 | 14 | 3 | 14 | 108 | even |
| 2297 | 19 | 47 | 308 | 16 | 10 | 34 | 66 | even |
| 3263 | 28 | 11 | 386 | 20 | 10 | 37 | 39 | odd |
| 5851 | 83 | 3 | 587 | 30 | 5 | 24 | 86 | even |
| 3925 | 37 | 137 | 451 | 23 | 4 | 22 | 174 | even |
| 5040 | 56 | 61 | 536 | 27 | 9 | 34 | 117 | odd |
| 1362 | 9 | 127 | 207 | 11 | 21 | 79 | 136 | even |
| 3514 | 32 | 11 | 415 | 21 | 11 | 42 | 43 | odd |
| 3815 | 37 | 27 | 447 | 23 | 5 | 24 | 64 | even |
| 5993 | 88 | 26 | 592 | 30 | 4 | 15 | 114 | even |
| 1403 | 10 | 39 | 213 | 11 | 19 | 79 | 49 | odd |
| 2719 | 23 | 46 | 345 | 18 | 7 | 38 | 69 | odd |
| 1529 | 11 | 56 | 228 | 12 | 18 | 65 | 67 | odd |
| 5431 | 71 | 12 | 571 | 29 | 9 | 41 | 83 | odd |
| 2843 | 24 | 52 | 356 | 18 | 10 | 45 | 76 | even |
| 4819 | 53 | 35 | 527 | 27 | 5 | 19 | 88 | even |
| 991 | 7 | 37 | 154 | 8 | 37 | 154 | 44 | even |
| 2971 | 26 | 39 | 368 | 19 | 5 | 22 | 65 | odd |
| 5553 | 75 | 2 | 577 | 29 | 4 | 20 | 77 | odd |
| 4890 | 54 | 44 | 530 | 27 | 5 | 20 | 98 | even |