DISCRETE PROJECT

(Cyber Security-T)

Name:
Muhammad Ismail Ramzan (20i-0941)
Muhammad Usman Shahid (20i-1797)
Syed Baqir Abbas (20i-0806)
Musaab Imran (20i-1794)

Submitted to: Ma'am Amna Irum

Contents

SECTION 01:

WEB SCRAPPING

Section 01:

Step 01:

First of all, we took three websites of three different universities as it was given in the instructions. All of then contained the noun "quality". In section 1 we had to do web scrapping to extract the data to have a comparison study among the different websites. In the comparative study we took in consideration nouns, adjectives, and verbs.

```python
from website1 import *
from website2 import *
from website3 import *
#Now i need access to the data of the website...
#print (first_website_total[0])
from matplotlib import pyplot
import numpy as np


# first_website_total,second_website_total & third_website_total
# 0 index contains the total number of Nouns
# 1 index contains total number of Verbs
# 2 Index contains total number of Ajection


# THIS Function takes 3 values and show the Graph with label for comparison
def comparison_graph_display(first_pof,second_pof,third_pof,labels):
    X = ["FAST","NUST","LUMS"]
    nouns = [first_pof,second_pof,third_pof]
    X_axis = np.arange(len(X))
    pyplot.bar(X_axis + 0.0, nouns, 0.4, label = labels)
    pyplot.xticks(X_axis, X)
    pyplot.xlabel(labels)
    pyplot.ylabel("Counting")
    pyplot.title("Comparison of Three Websites " + labels)
    pyplot.legend()
    pyplot.show()
# 0 index contains the total number of Nouns
# 1 index contains total number of Verbs
# 2 Index contains total number of Ajection
comparison_graph_display(first_website_total[0],second_website_total[0],third_website_total[0],"Nouns")
comparison_graph_display(first_website_total[1],second_website_total[1],third_website_total[1],"Verbs")
comparison_graph_display(first_website_total[2],second_website_total[2],third_website_total[2],"Adjective")
```

When all of the text from the different websites was extracted then we moved on to the next step of extracting the nouns, adjectives, and verbs.

Website 01:

```python
# website1.py > ...
1   from header import *
2
3   first_website = ["http://nu.edu.pk/","http://isb.nu.edu.pk/","http://khi.nu.edu.pk/",
4                    "http://lhr.nu.edu.pk/","http://pwr.nu.edu.pk/"]
5   # Calling the function to go and get the web data
6   #############
7   first_website_text = [] # Contains the content of the website |
8   first_website_nouns = [] # Contains the number of nouns in website
9   first_website_verbs = []  # Contains the verbs of the website
10  first_website_adj = []  # Contains the adjective of the website
11  ################## Go and show the Graph with Subdomains #################
12  show_subdomain_graph(first_website_text,first_website_nouns,first_website_verbs,first_website_adj,first_website)
13  ################# Go and count the total number of Parts of speech in doamin ########
14  first_website_total = []
15  # 0 index contains the total number of Nouns
16  # 1 index contains total number of Verbs
17  # 2 Index contains total number of Ajection
18  first_website_total.append(count_total(first_website_nouns))
19  first_website_total.append(count_total(first_website_verbs))
20  first_website_total.append(count_total(first_website_adj))
21  ################# Show the Final Graph of the Website ##############3
22  print_graph(first_website_total[1],first_website_total[2],first_website_total[0],"Total For this Domain")
23
24
```

Website 02:

```python
# website2.py > ...
1   from header import *
2
3   second_website = ["https://nust.edu.pk/","https://nust.edu.pk/admissions/",
4                     "https://nust.edu.pk/academics/","https://nio.nust.edu.pk/",
5                     "https://campuslife.nust.edu.pk/"]
6   # Calling the function to go and get the web data
7   #############
8   second_website_text = [] # Contains the content of the website
9   second_website_nouns = [] # Contains the number of nouns in website
10  second_website_verbs = []  # Contains the verbs of the website
11  second_website_adj = []  # Contains the adjective of the website
12  ################# Go and show the Graph with Subdomains ################
13  show_subdomain_graph(second_website_text,second_website_nouns,second_website_verbs,
14                       second_website_adj,second_website)
15  ################# Go and count the total number of Parts of speech in doamin ########
16  second_website_total = []
17  # 0 index contains the total number of Nouns
18  # 1 index contains total number of Verbs
19  # 2 Index contains total number of Ajection
20  second_website_total.append(count_total(second_website_nouns))
21  second_website_total.append(count_total(second_website_verbs))
22  second_website_total.append(count_total(second_website_adj))
23  ################# Show the Final Graph of the Website ##############3
24  print_graph(second_website_total[1],second_website_total[2],second_website_total[0],"Total For this Domain")
25
26
```

Website 03:

```python
from header import *

third_website = ["https://lums.edu.pk/","https://admission.lums.edu.pk/",
                 "https://lums.edu.pk/news","https://lums.edu.pk/student-noticeboard",
                 "https://admission.lums.edu.pk/admission-policy"]
# Calling the function to go and get the web data
#############
third_website_text = [] # Contains the content of the website
third_website_nouns = [] # Contains the number of nouns in website
third_website_verbs = []  # Contains the verbs of the website
third_website_adj = []  # Contains the adjective of the website
################# Go and show the Graph with Subdomains ################
show_subdomain_graph(third_website_text,third_website_nouns,third_website_verbs,third_website_adj,third_website)
################ Go and count the total number of Parts of speech in doamin ########
third_website_total = []
# 0 index contains the total number of Nouns
# 1 index contains total number of Verbs
# 2 Index contains total number of Ajection
third_website_total.append(count_total(third_website_nouns))
third_website_total.append(count_total(third_website_verbs))
third_website_total.append(count_total(third_website_adj))
################ Show the Final Graph of the Website ###############3
print_graph(third_website_total[1],third_website_total[2],third_website_total[0],"Total For this Domain")
```

Sample of extracted text from that particular website using the above-mentioned code.

. hrs.TitleCr. hrs.Philosophy of Management Science3Quantitative Methods in Management Research 3Research and Theory in Management3Qualitative Methods in Management Research3Note: Deficiency courses may be recommended by FSM-DAC for selected candidates on case to case basis.Outlines of Core Courses (These are broad guidelines only. Instructors may vary emphases on individual topics and themes and bring additional topics that they see fit)Philosophy of Management Science (MG601): It is said that management is governing and leading without recourse to political power or religious authority. This advanced level course will help students to realize and appreciate the nature of management field to a greater maturity and sophistication by interchangeably addressing the philosophical aspects of management and managerial aspects of philosophy. Accordingly, the process that will be used will start with, and put major focus on, developing an understanding of the evolution of management and organizational philosophy in a chronological order by classifying the work of central theorists and authors from a perspective that clarifies their relevance with the mainstream theory building in the field of management. The idea is to dissect the body of knowledge in management science \xe2\x80\x93 i.e., perspectives such as classical and neo-classical organization theory, OB perspective, structural organization theory, systems theory and organizational economics, power and politics, organizational culture and environment, and postmodernism, etc. \xe2\x80\x93 and interpret it as a continuous effort to analyzing, predicting, controlling and governing human action.Research and Theory in Management (MG604): A theory is said to be an ordered set of assertions about a generic set of behavior that hold true for a wide range of specific instances. But where do new theories come from? How theories are built starting from a very basic research question? And can theories be good or bad since once popularized they not only explain behavior but influence behavior in a certain direction. Finally, what constitutes a \xe2\x80\x98weak\xe2\x80\x99 or a \xe2\x80\x98strong\xe2\x80\x99 theory? The aim of this course is to develop a critical appreciation ofthe role of research in the ory development in Management as well as introducing doctoral students to the art and science of theory construction and model building. Topics include examining the nature of being (ontology) and knowing (epistemology), exploring the basic structure of argument and evidence, developing a pertinent and interesting research question, defining constructs, imagining relationships and processes that link constructs, and linking and combining constructs in an innovative way to develop new theories or build further upon existing behavioral models. This course demands intensive preparation on part of students in terms of readings prior to the lectures so that lively discussions can be generated in the classroom, research exercises tailored to each student\xe2\x80\x99s research topic can be accomplished, and a term paper be completed.Qualitative Methods (MG 589): This course will introduce a range of qualitative research methods and the means of assessing the key assumptions underlying the selection and use these methods. Building on the fundamental questions about the nature of knowledge and whether "true" knowledge or reality can exist, the discussions will provide a deep understanding of the advantages and challenges in carrying out qualitative research. Examples of topics covered include interviews, focus groups, ethnography, documentary and textual analysis, discourse analysis, language and conversation analysis, case-studies, participatory action research and grounded theory. Categorization and coding of data will be covered in detail. The course will include labs to introduce software for qualitative data analysis. The course will also cover the growing significance of qualitative research techniques across the modern and postmodern era. Further, scholars will be encouraged to analyze and understand various approaches to writing qualitative papers.Quantitative Methods (MG 548): This course will introduce a range of advance quantitative research technique and develop an understanding of the potential and limitations of statistical analysis. Building on the positivist paradigm, the course will briefly revise concepts of probability, sampling, hypothesis development, developing constructs, associations and basic analyses for differences between groups. Examples of main topics covered include various types of regression, discriminant analysis, MANOVA, analysis of covariance, principal component and factor analysis, cluster analysis, structural equation modelling, survival/failure analysis and time series analysis. The course will include labs to introduce software for quantitative data analysis. Scholars will be encouraged to analyze and understand articles which have used a variety of quantitative methods, developing the capacity to choose advance techniques suitable for their research questions. More Details PhD (Software Engineeri

SECTION 02:
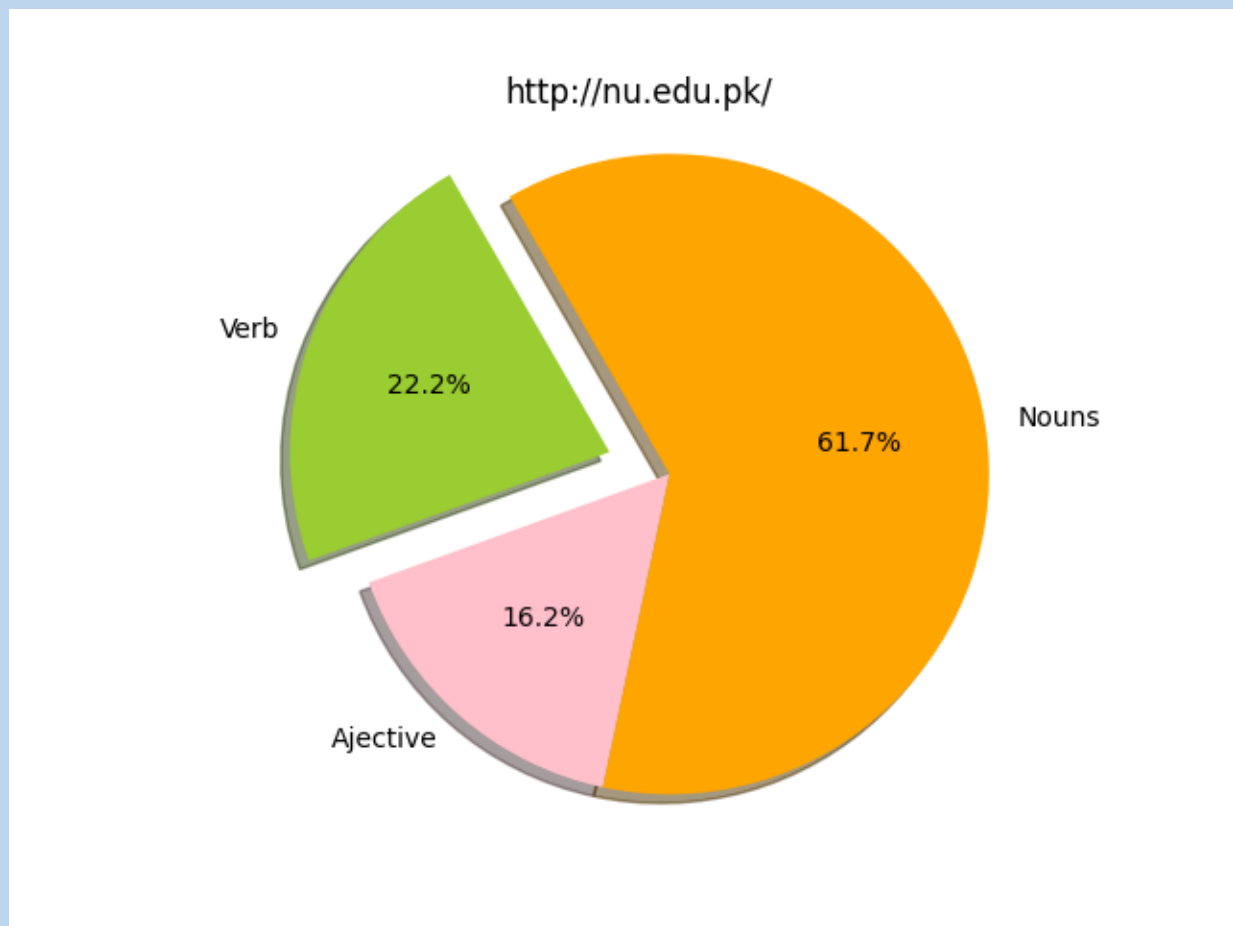# EXTRACTION OF:
1. NOUNS
2. VERBS
3. ADJECTIVES

Section 02:

Website 01 (NUCES):

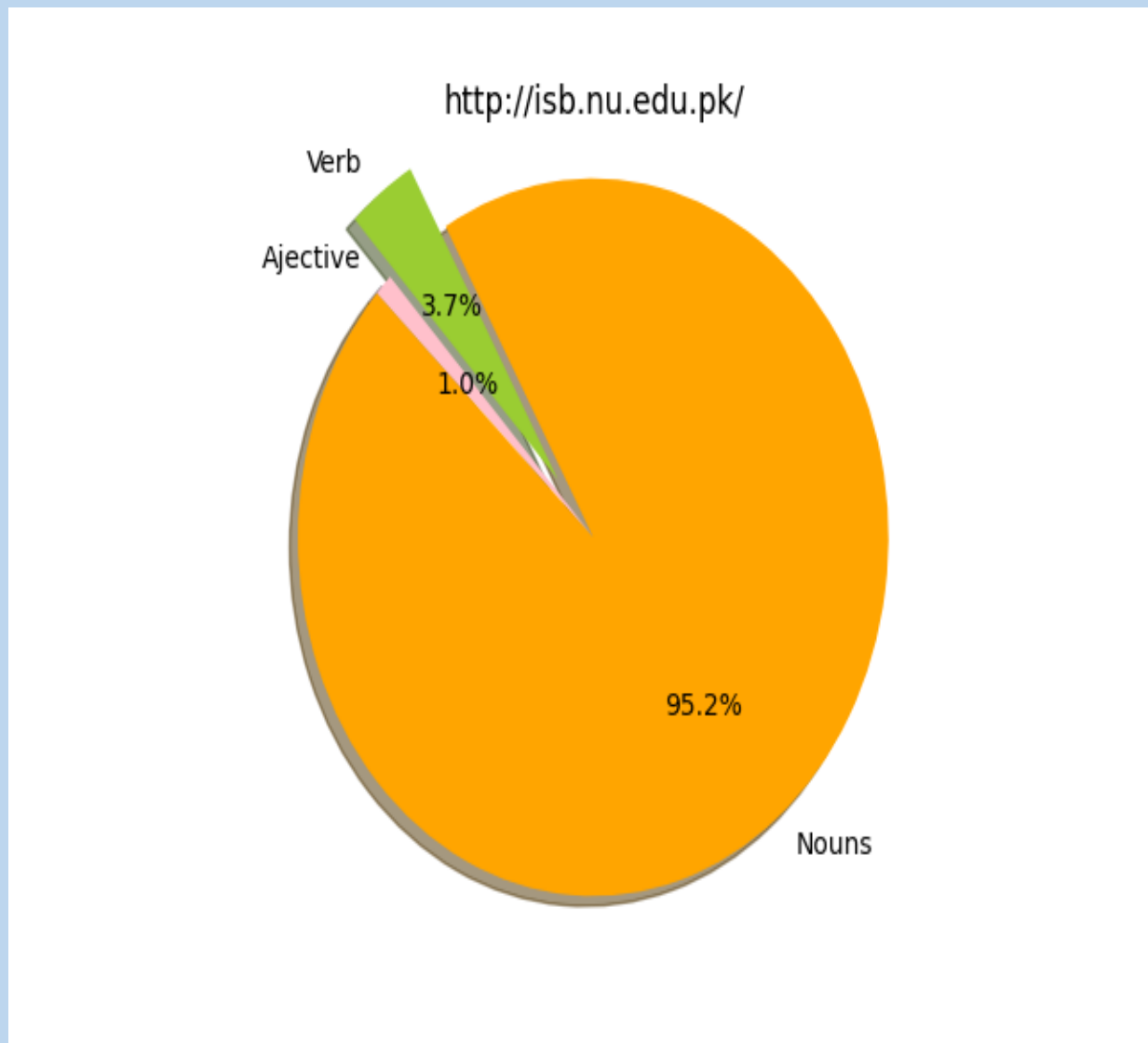**Web Page 01:**

Verbs: 1299

Nouns: 3615

Adjectives: 947

http://nu.edu.pk/

Web Page 02:

Verbs: 22

Nouns: 560

Adjectives: 6

http://isb.nu.edu.pk/

Verb

Ajective

3.7%

1.0%

95.2%

Nouns

Web page 03:

Verbs: 104

Nouns: 492

Adjectives: 74

http://khi.nu.edu.pk/

Verb 15.5%

Ajective 11.0%
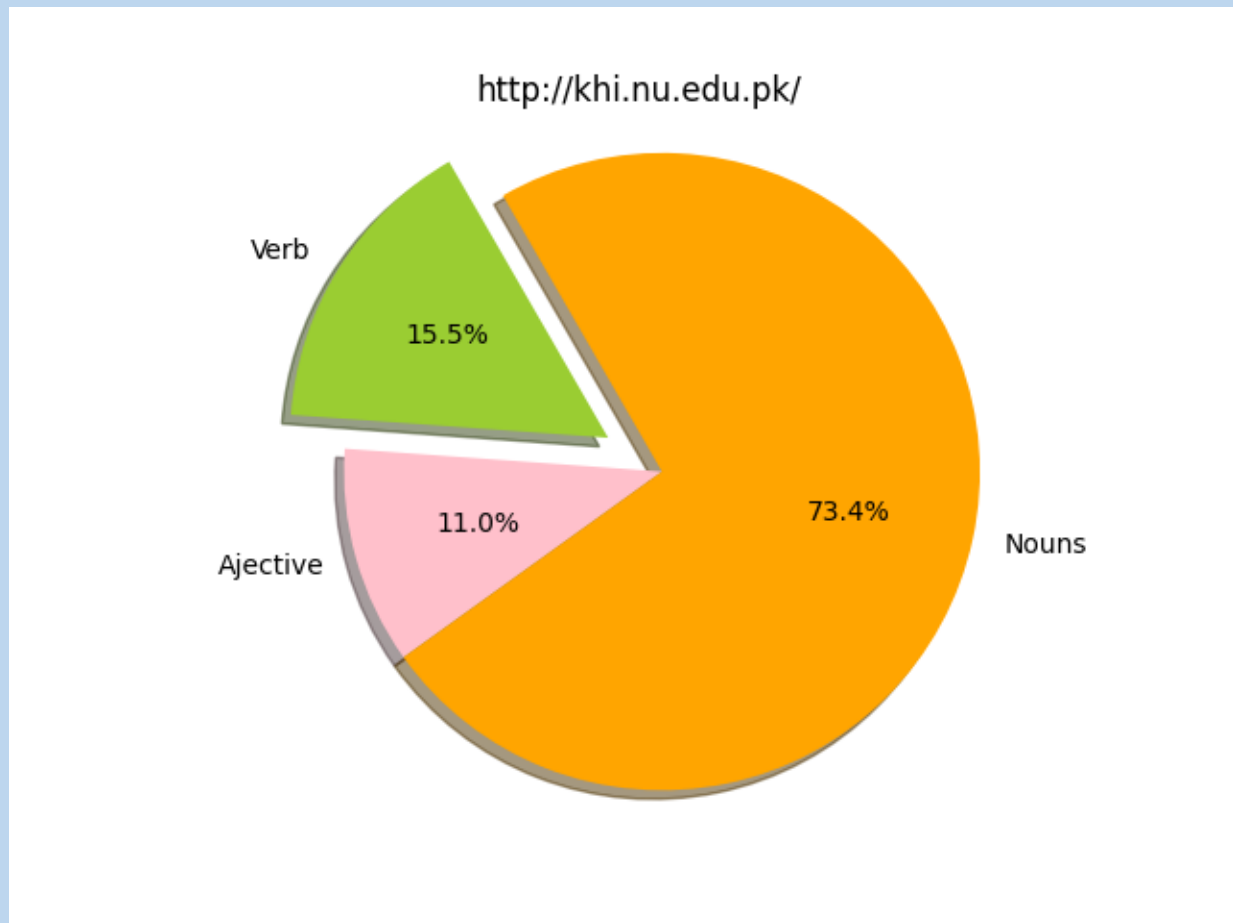
Nouns 73.4%

Web page 04:

Verbs: 52

Nouns: 262

Adjectives: 40

http://lhr.nu.edu.pk/

Verb

11.5%

Ajective

8.8%

79.7%
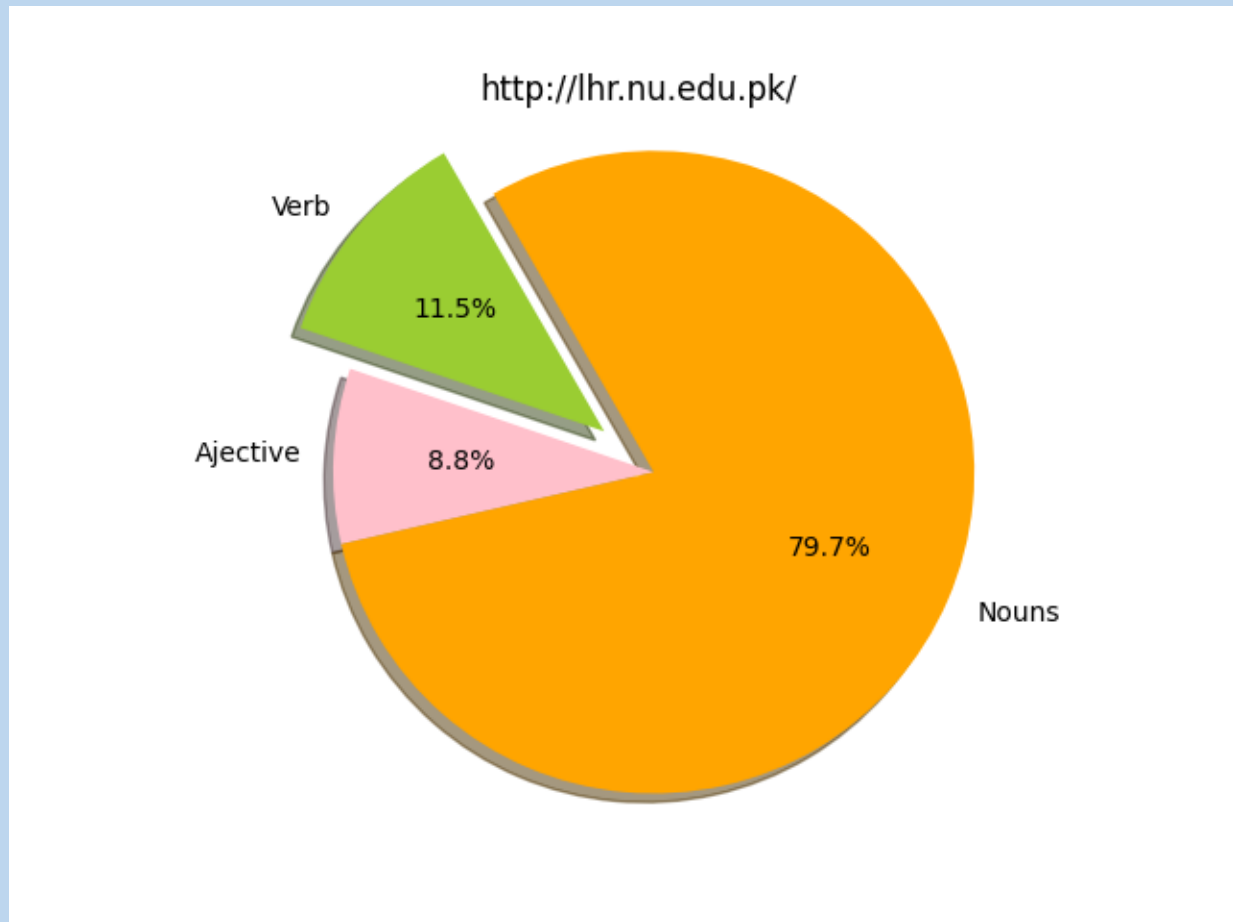
Nouns

Web page 05:

Verbs: 303

Nouns: 1110

Adjectives: 146

http://pwr.nu.edu.pk/

Verb 19.4%

Ajective 9.4%
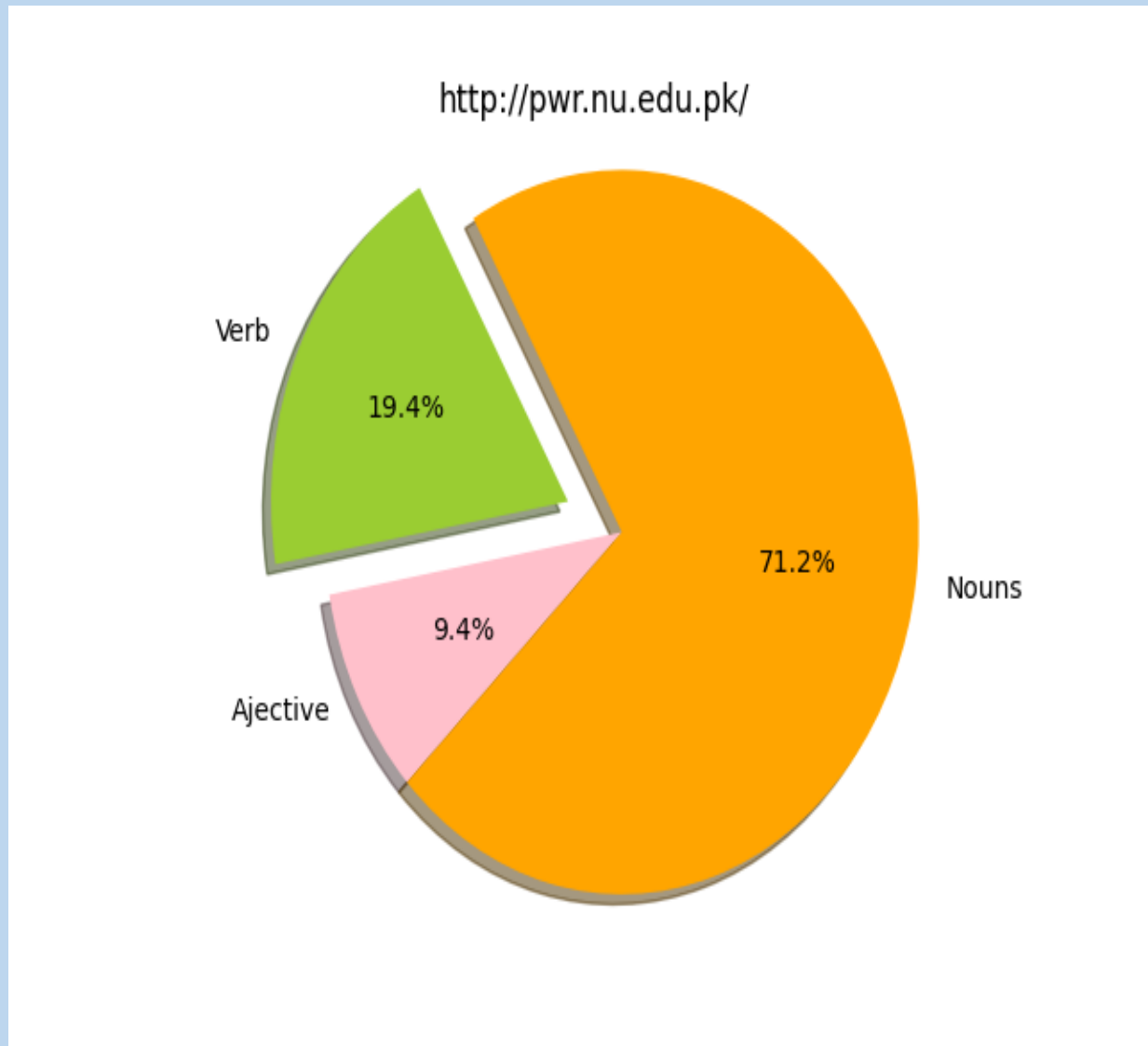
Nouns 71.2%

Total number of nouns, adjectives, and verbs of FAST website:

Nouns = 6039

Verbs = 1780

Adjectives = 1231

Total For this Domain

Verb 19.5%

Nouns 67.2%

Ajective 13.3%
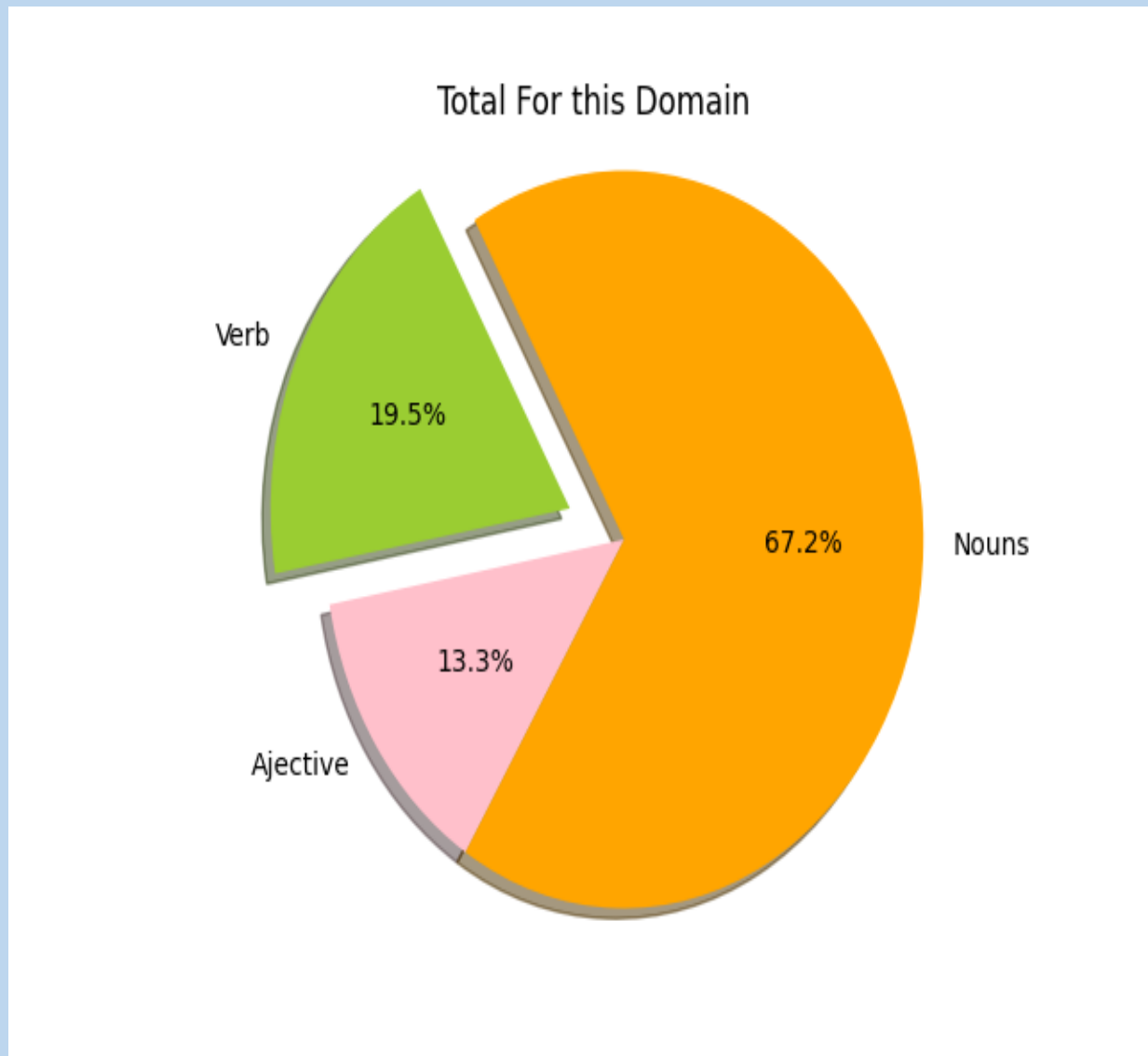
Website 02 (NUST):

Web Page 01:

Verbs: 46

Nouns: 709

Adjectives: 37

https://nust.edu.pk/

- Verb — 5.8%
- Ajective — 4.7%
- Nouns — 89.5%
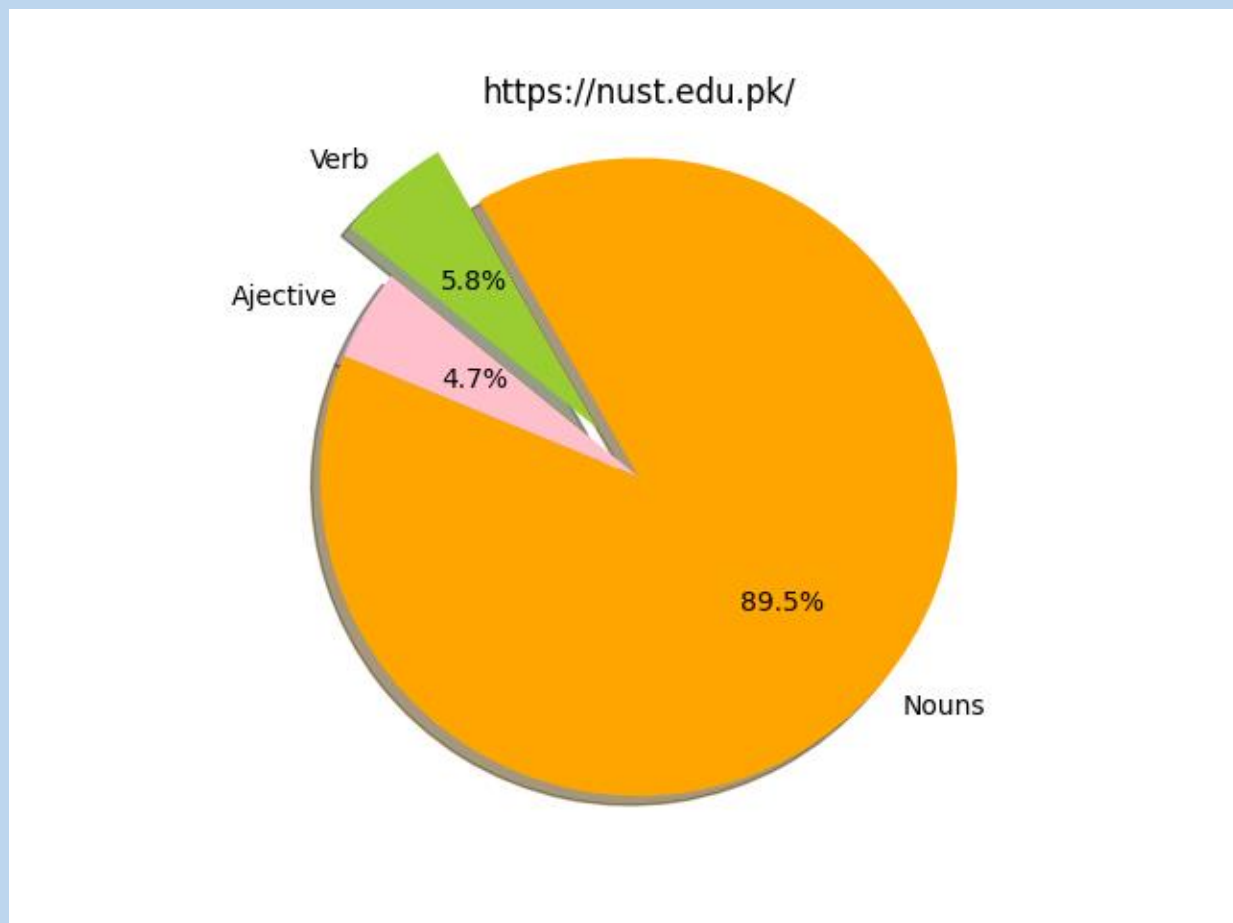
Web Page 02:

Verbs: 8

Nouns: 535

Adjectives: 11

https://nust.edu.pk/admissions/

Verb

Ajective

1.4%

2.0%

96.6%
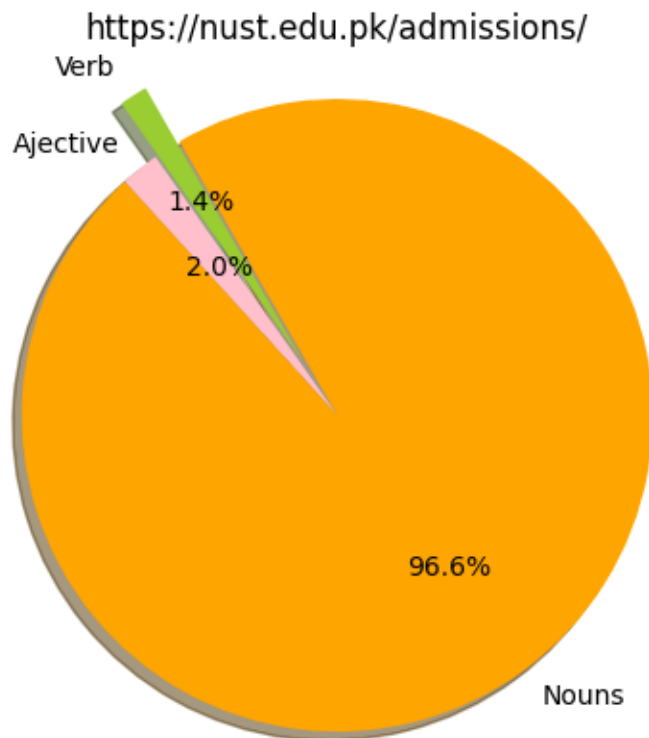
Nouns

Web Page 03:

Verbs: 2

Nouns: 202

Adjectives: 1

https://nust.edu.pk/academics/

Verb

Ajective

1.0%

0.5%

98.5%
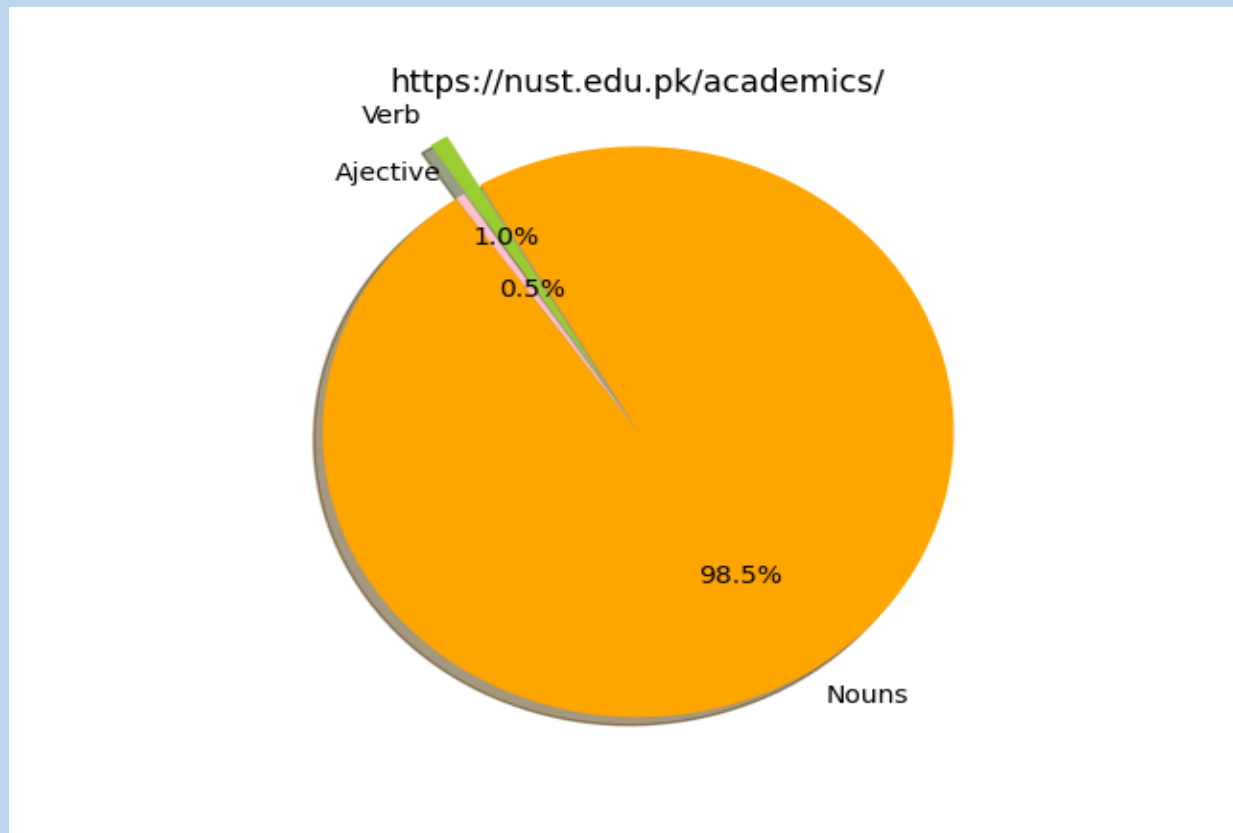
Nouns

Web Page 04:

Verbs: 18

Nouns: 234

Adjectives: 18

https://nio.nust.edu.pk/

Verb
6.7%

Ajective
6.7%

86.7%
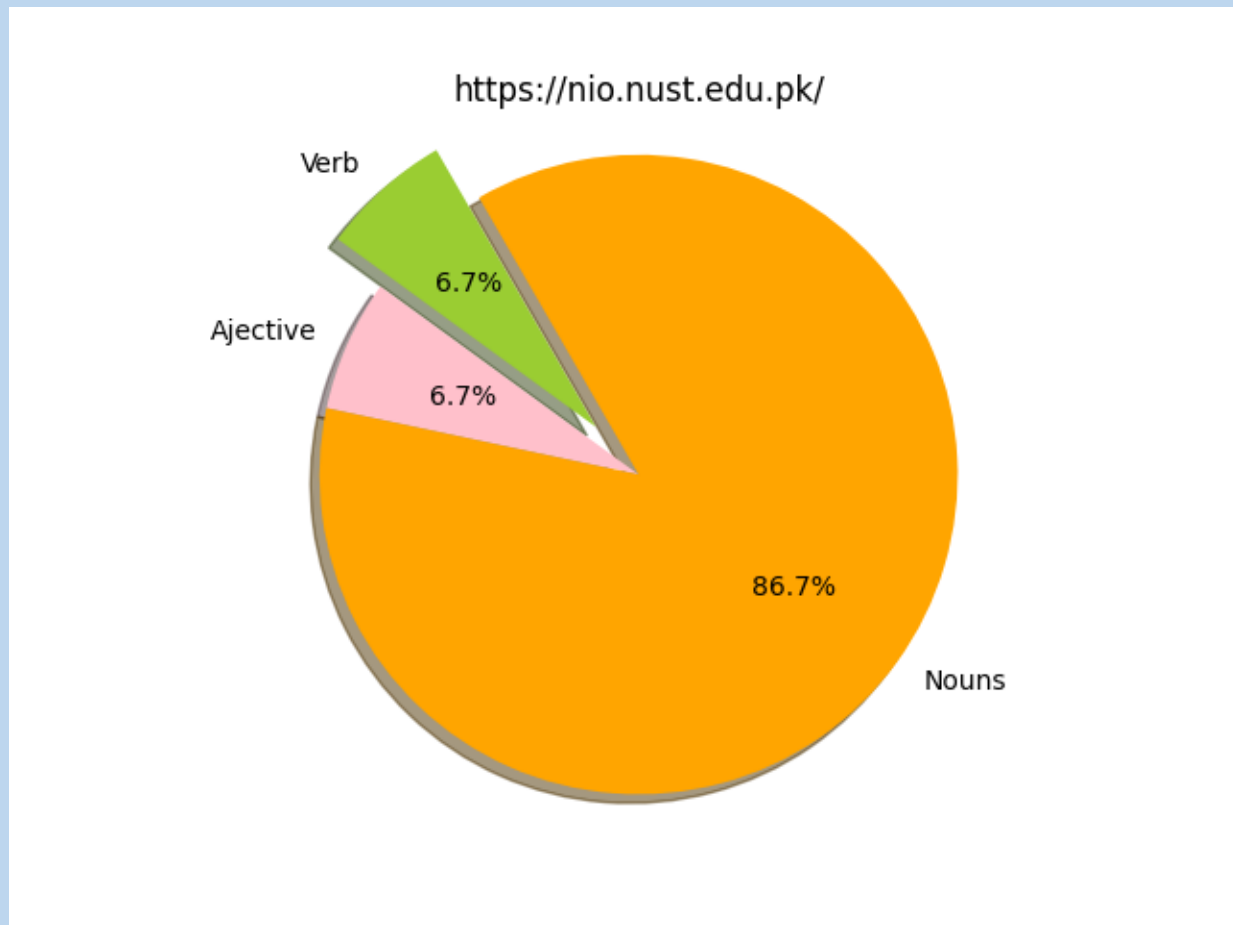
Nouns

Web Page 05:

Verbs: 21

Nouns: 301

Adjectives: 18

https://campuslife.nust.edu.pk/

Verb 6.2%

Ajective 5.3%
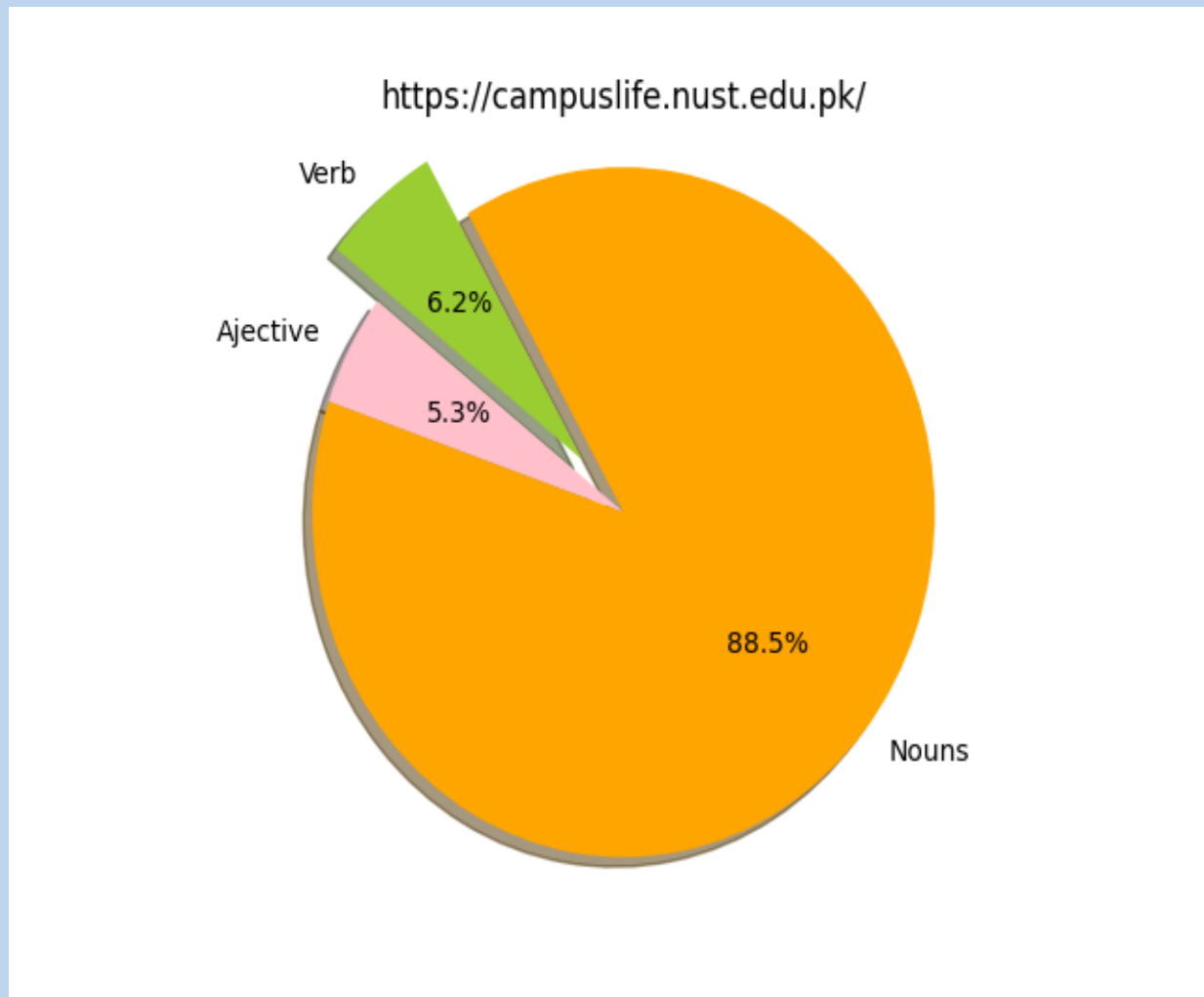
Nouns 88.5%

Total number of nouns, adjectives, and verbs of NUST website:

Nouns = 1719

Verbs = 95

Adjectives = 85
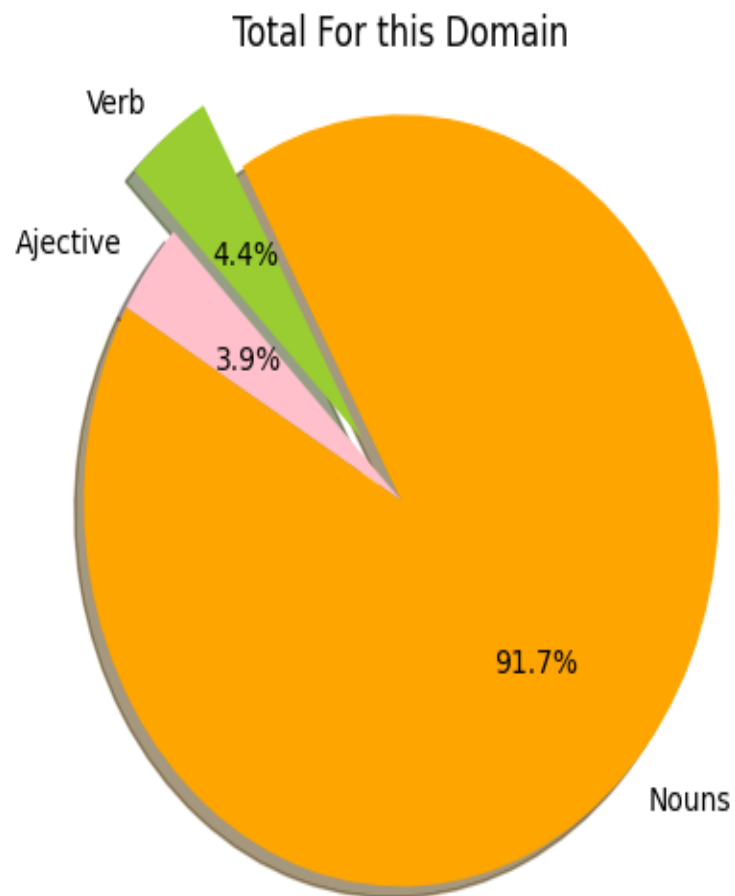
Website 03 (LUMS):

Web Page 01:

Verbs: 87

Nouns: 841

Adjectives: 68
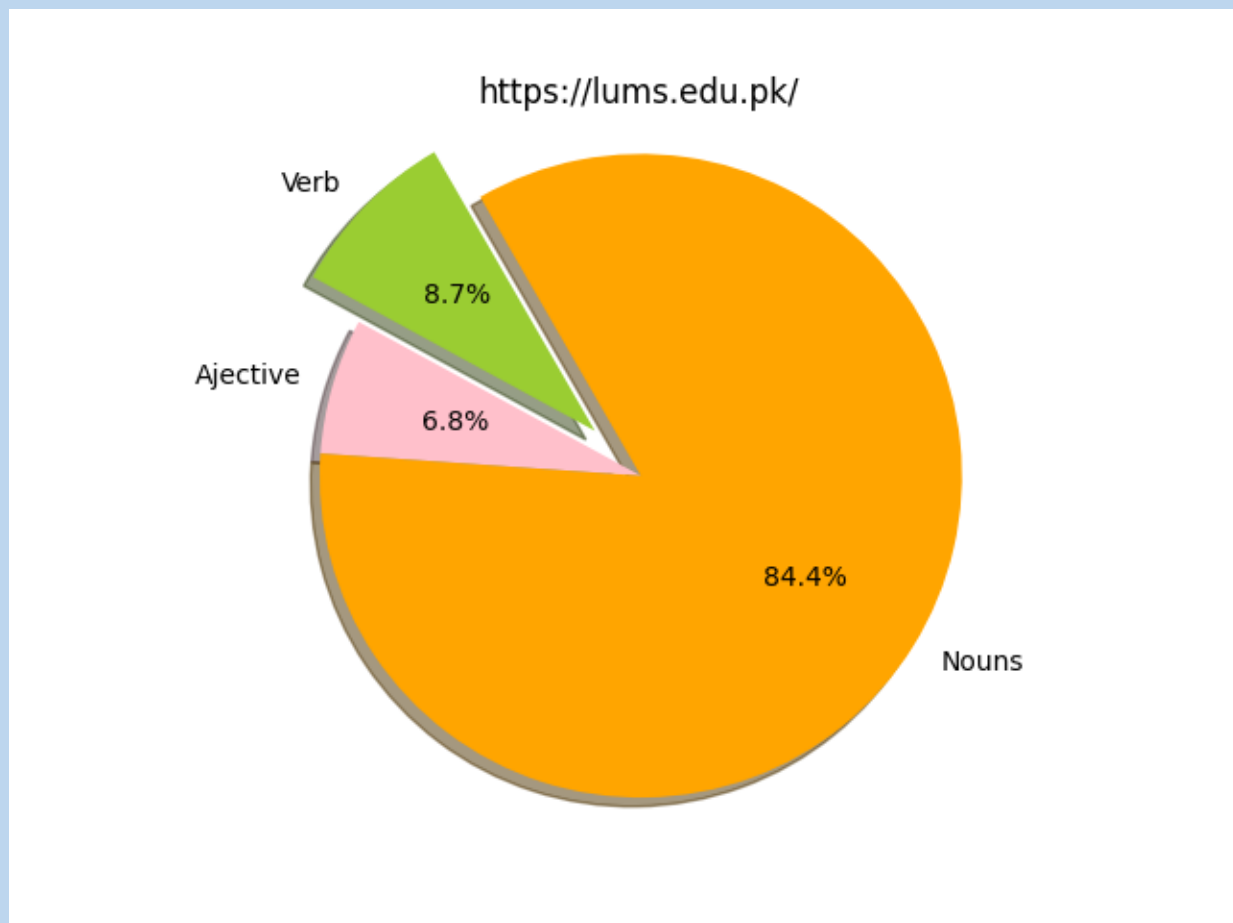
Web Page 02:
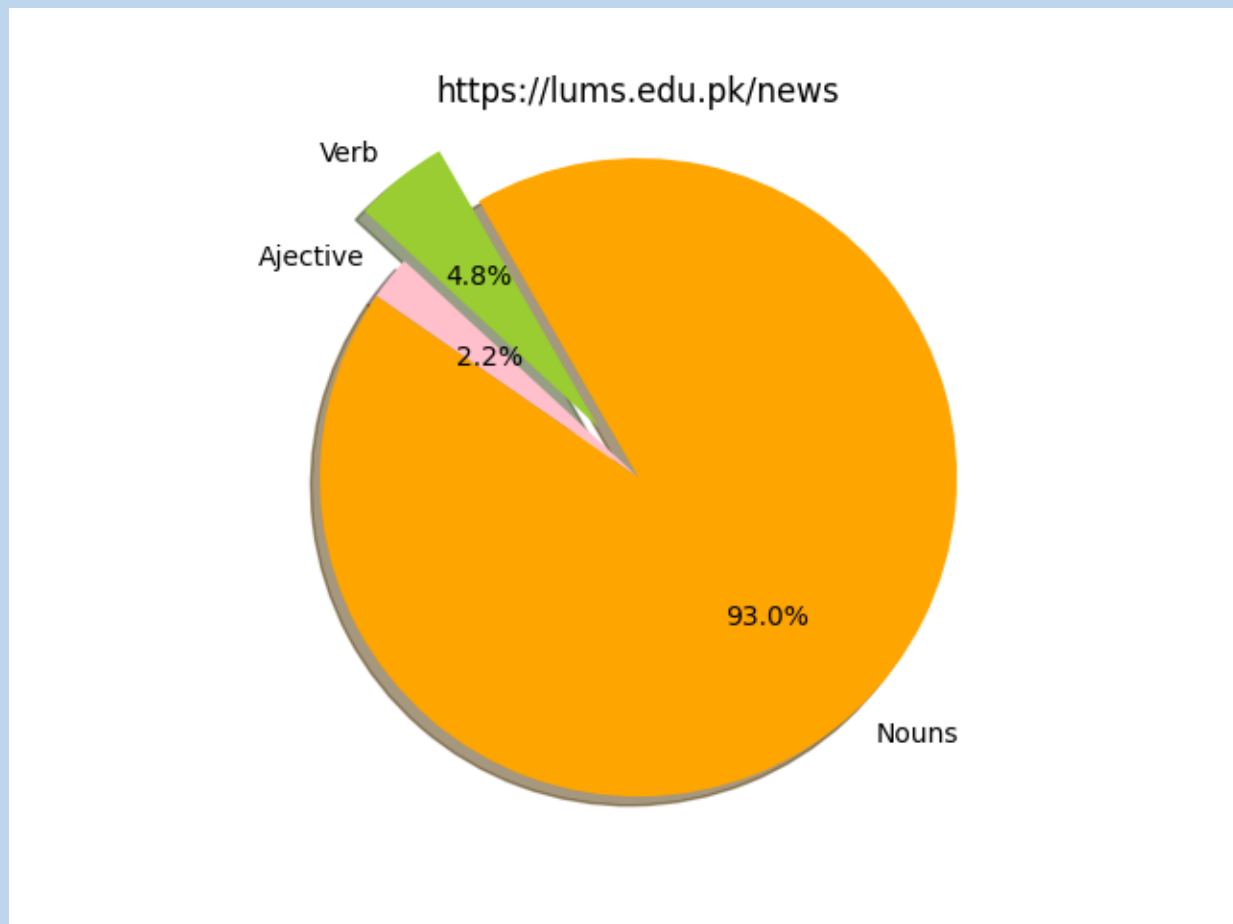
Verbs: 44

Nouns: 652

Adjectives: 27

https://admission.lums.edu.pk/

Verb

Ajective

6.1%

3.7%

90.2%
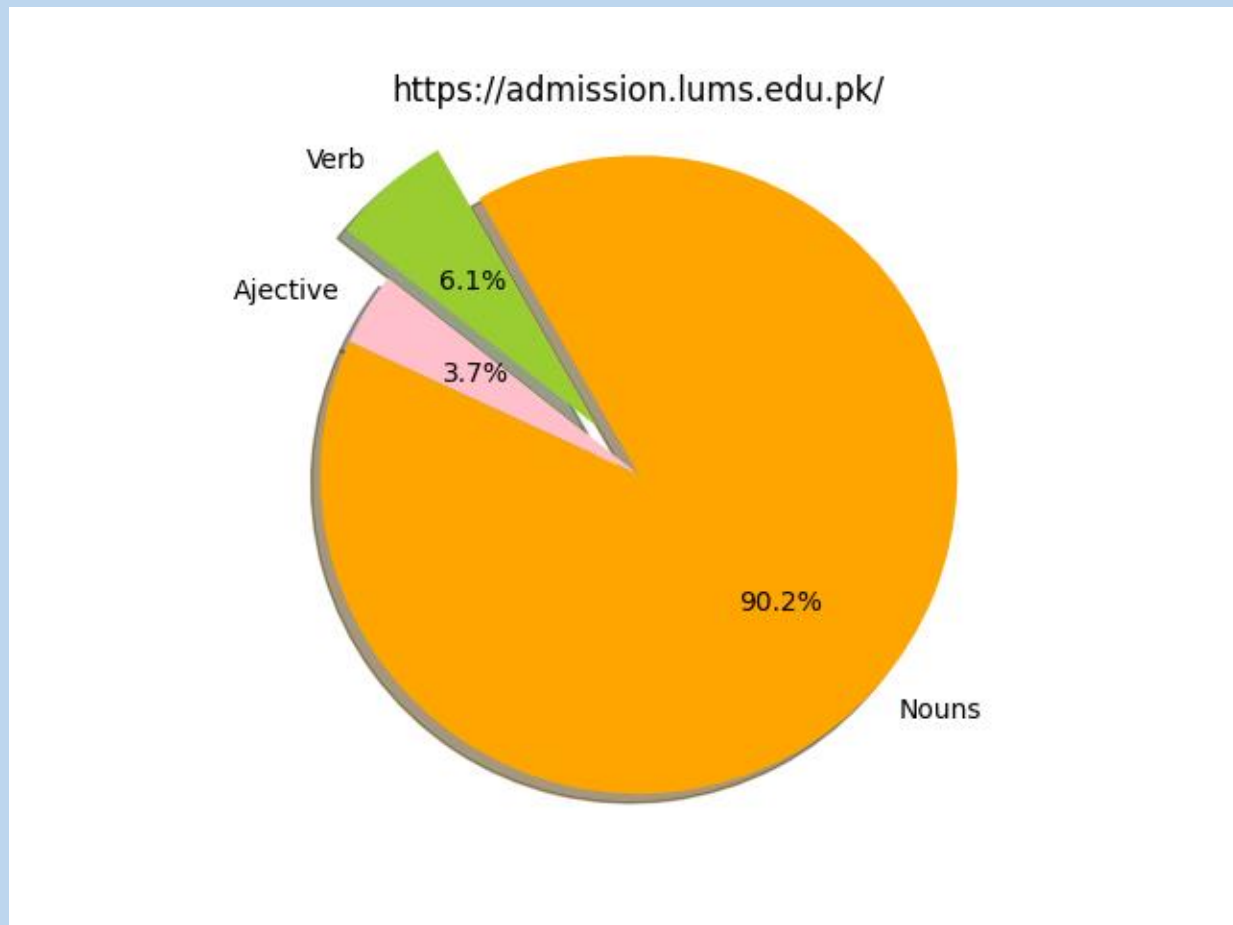
Nouns

Web Page 03:

Verbs: 30

Nouns: 581

Adjectives: 14

https://lums.edu.pk/news

Verb

Ajective

4.8%

2.2%

93.0%

Nouns

Web Page 04:

Verbs: 2389

Nouns: 4672

Adjectives: 1163

https://lums.edu.pk/student-noticeboard

Verb 29.0%

Nouns 56.8%

Ajective 14.1%
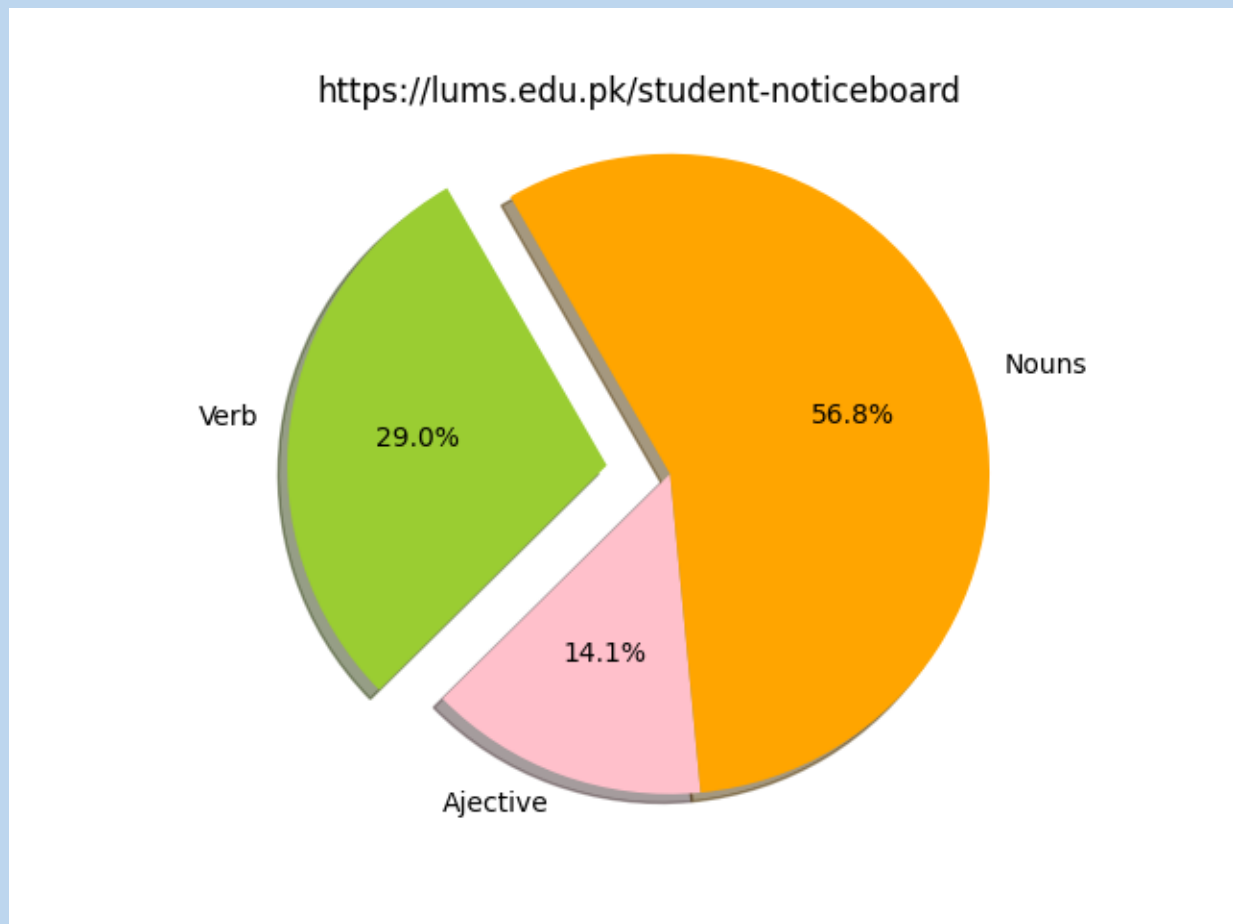
Web Page 05:

Verbs: 51

Nouns: 525

Adjectives: 42

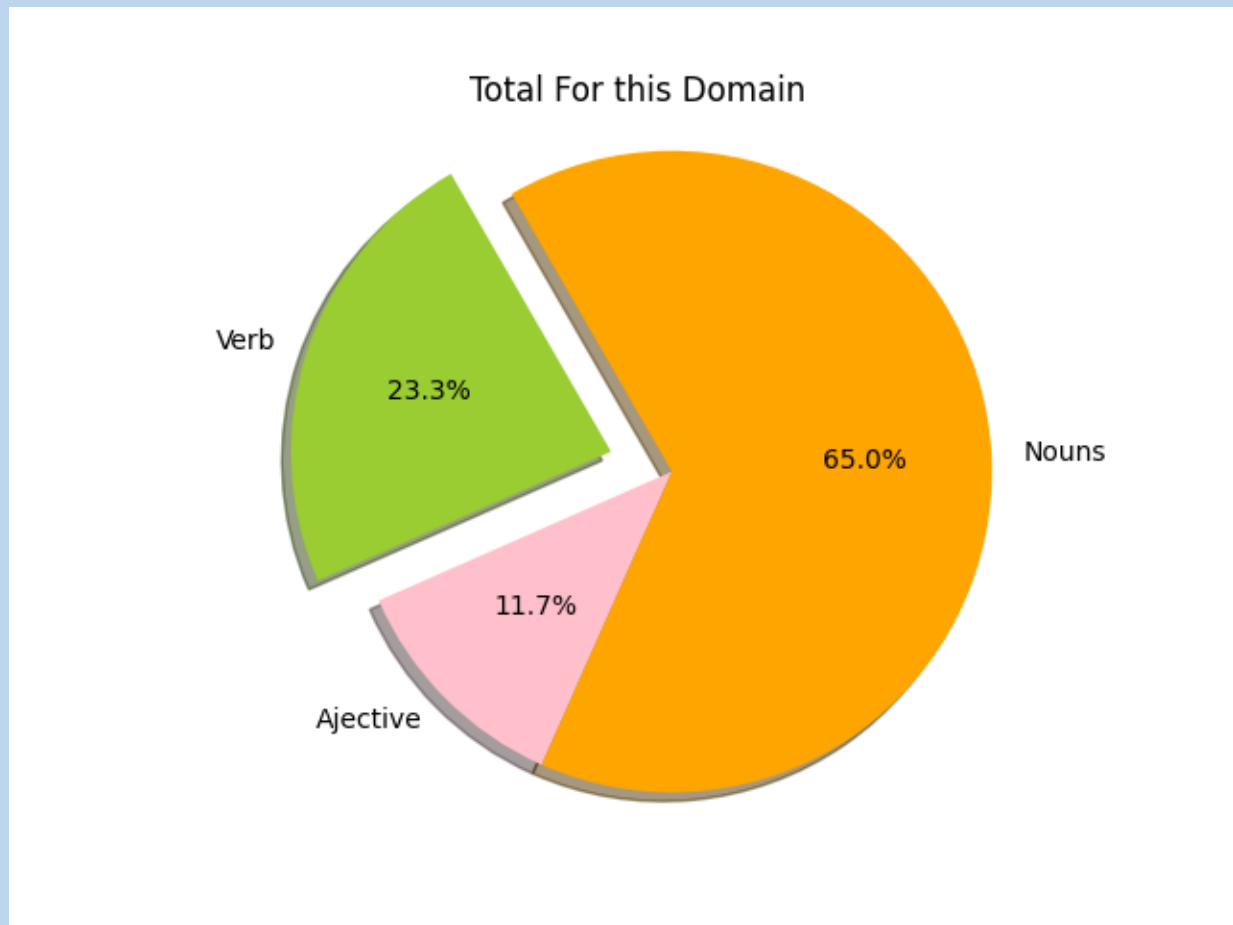https://admission.lums.edu.pk/admission-policy

Verb 8.3%

Ajective 6.8%

Nouns 85.0%
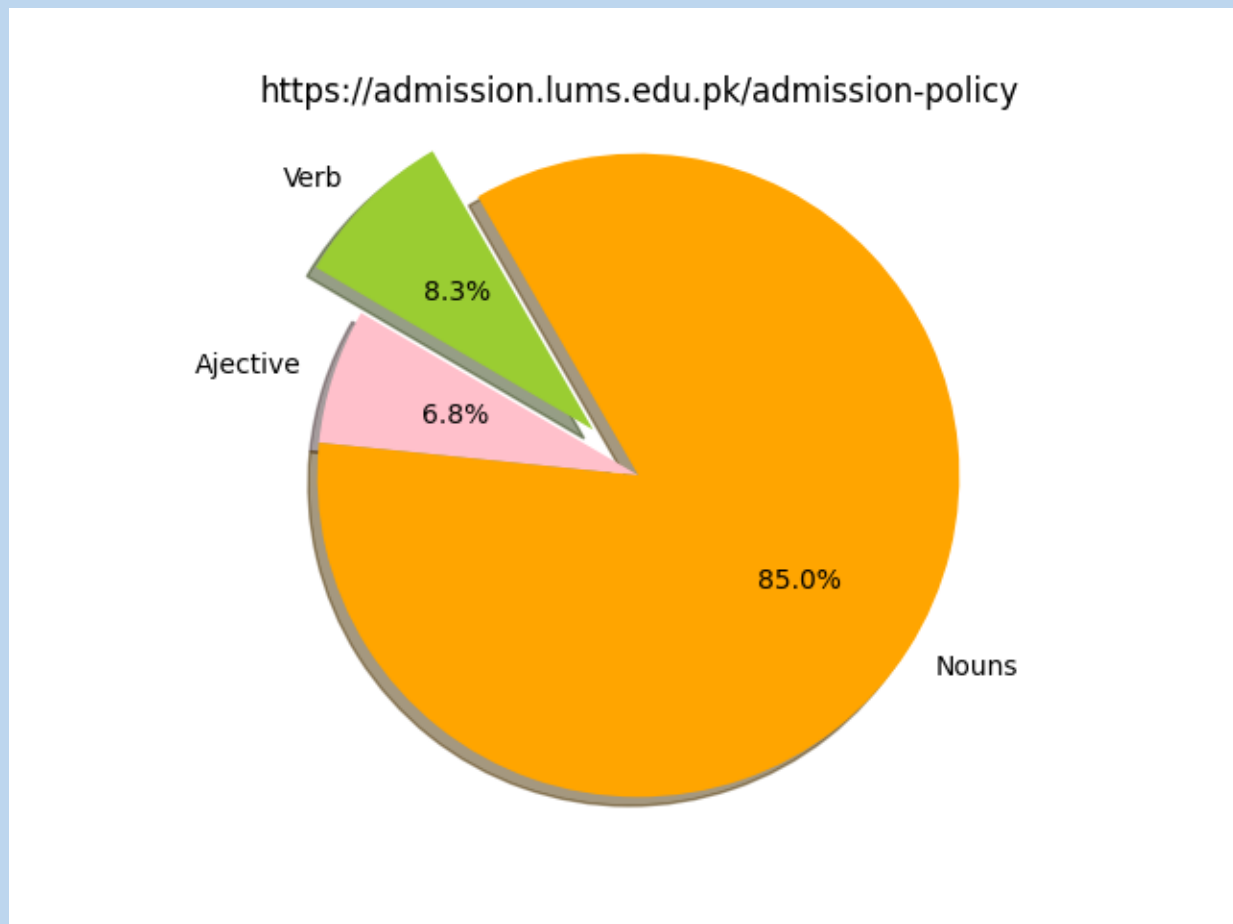
Total number of nouns, adjectives, and verbs of LUMS website:

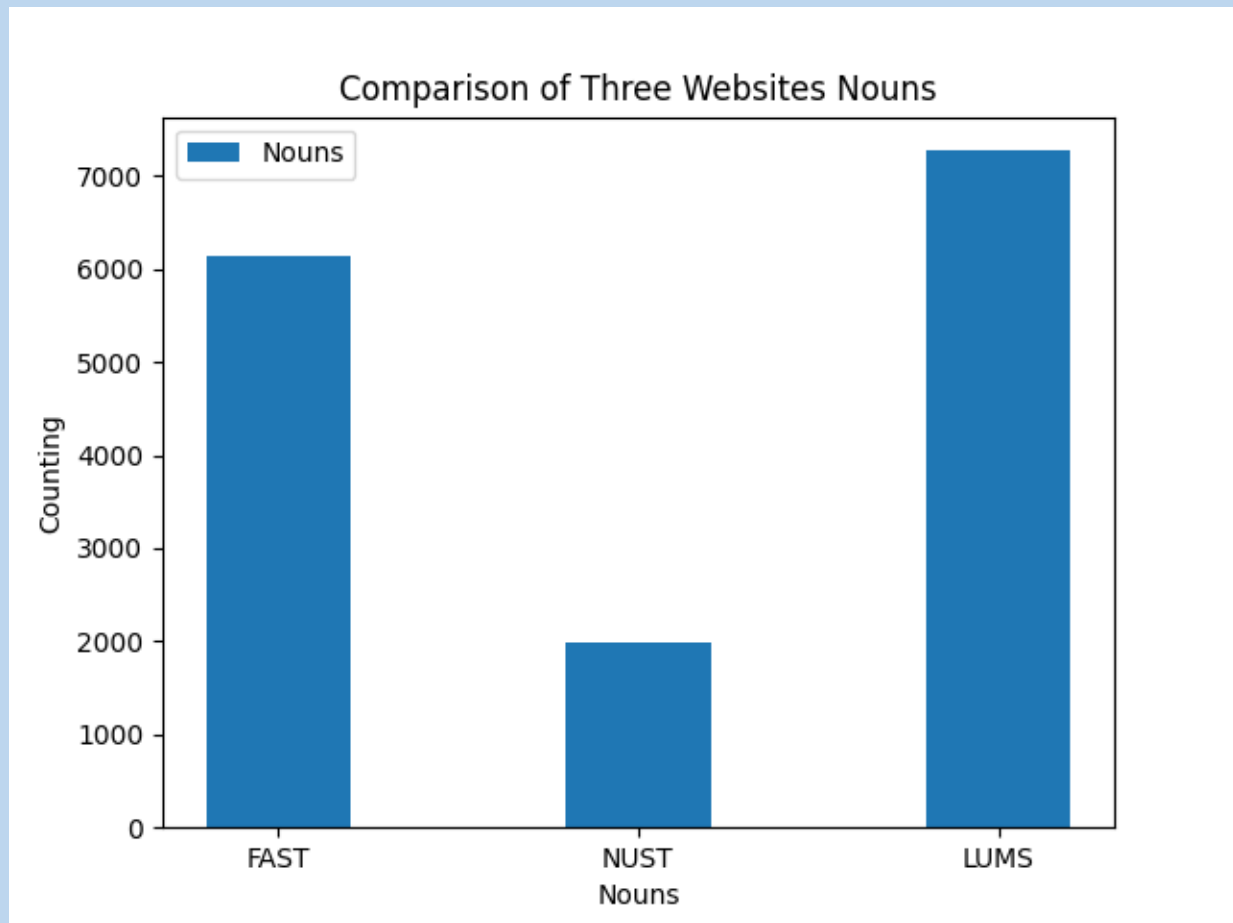Nouns = 7271

Verbs = 2601

Adjectives = 1314

Total For this Domain

COMPARISON OF VERBS, ADJECTIVES AND NOUNS AMONG THE THREE WEBSITES
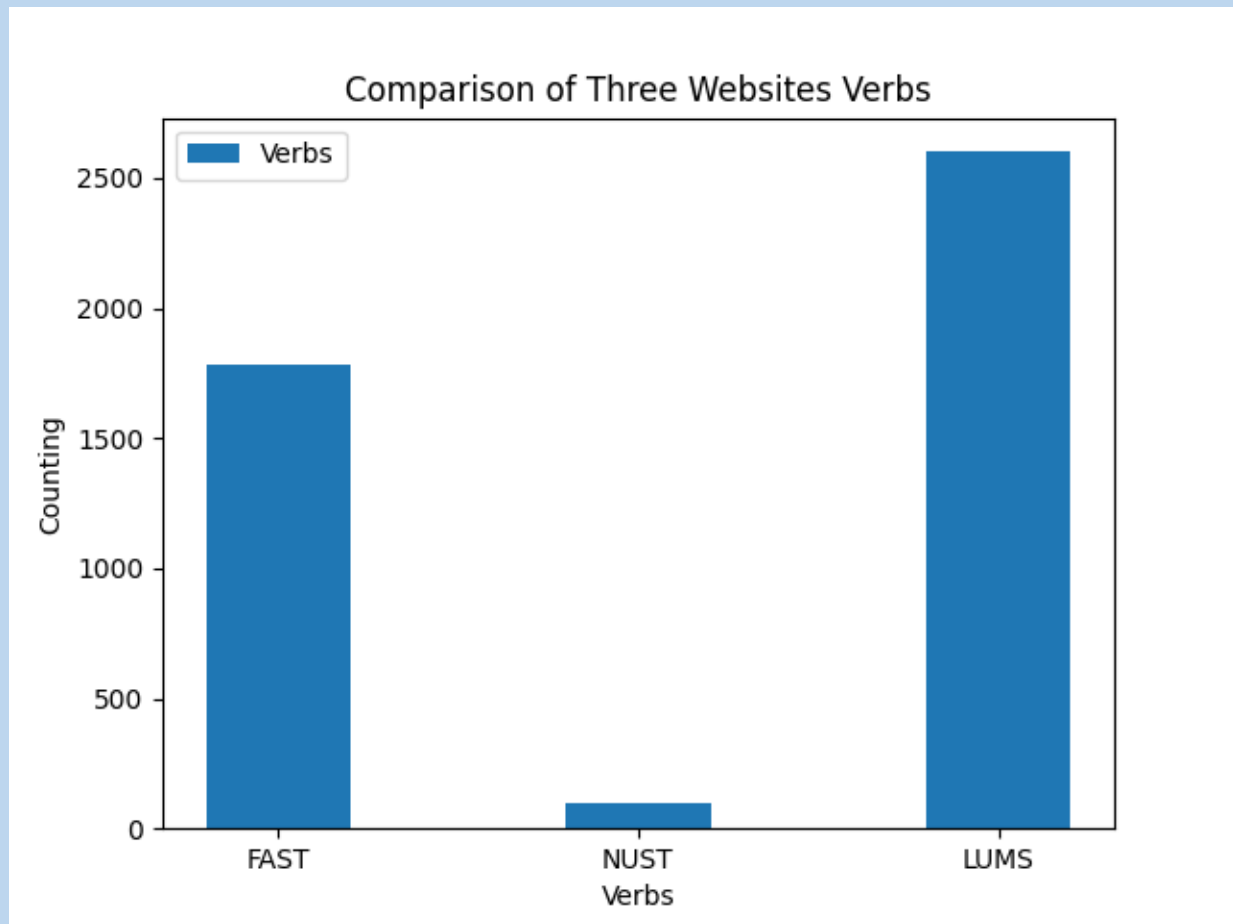
COMPARISON

Comparison of nouns:

From the noun comparison bar chart, we can see that the total number of nouns for all the websites were plotted in the chart. The 5 web pages that were selected from each website contained many nouns. So, the number of nouns (7271) in LUMS university website are more than NUST and FAST university website. NUST university website contained the least number of nouns.

Comparison of verbs:

When we compared verbs from all three website, we still can see that LUMS had the highest number of verbs (2601) as compared to the rest two university websites.

Comparison of Adjectives:
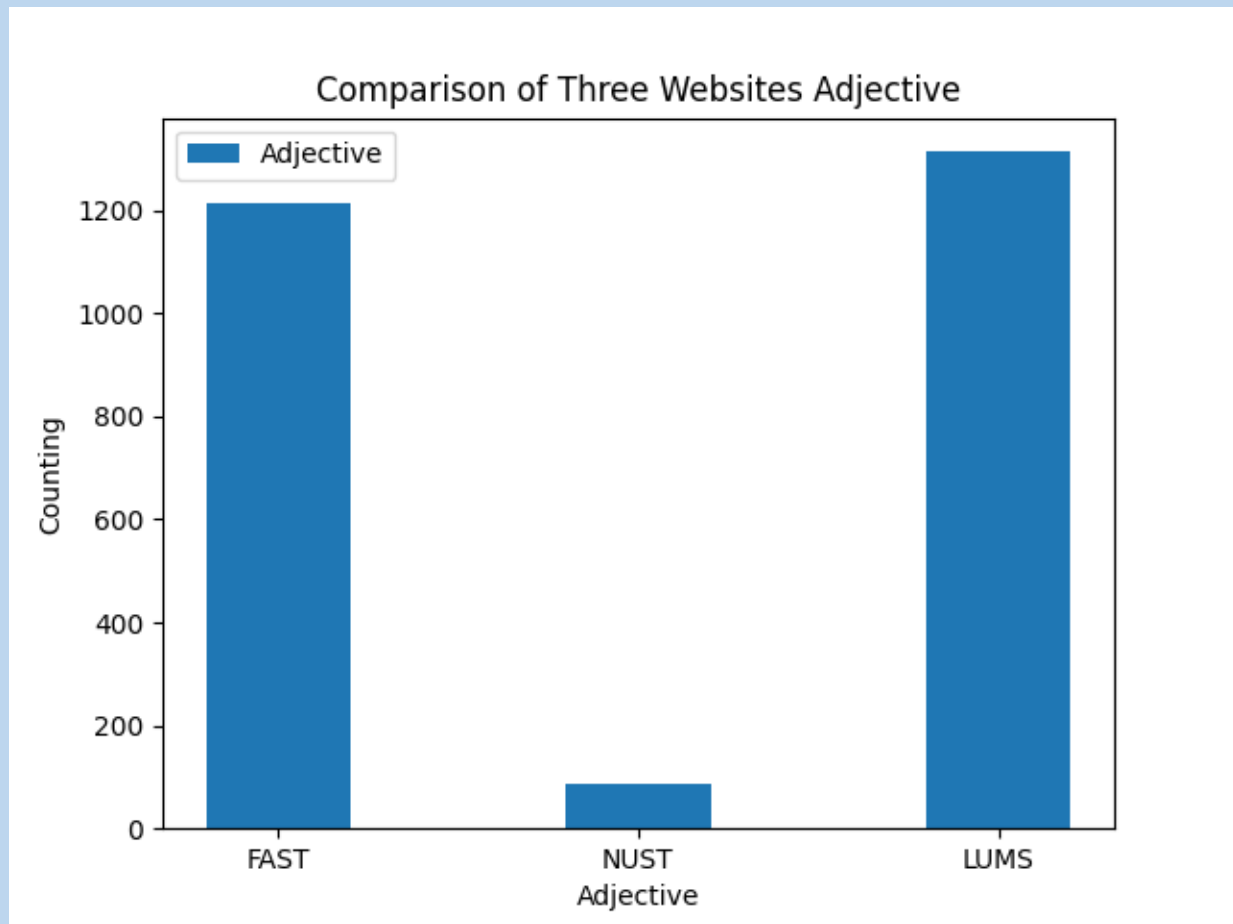
When we compared adjectives from all three website, we still can see that LUMS had the highest number of adjectives (1314) as compared to the rest two university websites.

Data from all three websites:

Top 10 nouns, adjectives, and verbs with their weights of first university website (NUCES).

```
TOP 10 Nouns of First Website

[['courses', 111], ['research', 74], ['student', 59], ['PhD', 57], ['program', 56], ['students', 46], ['degree', 39], ['knowledge', 39], ['credit', 35], ['management', 32]]

TOP 10 Verbs First Website

[['have', 54], ['are', 36], ['including', 28], ['been', 24], ['specified', 23], ['Passed', 22], ['passed', 19], ['computing', 17], ['conduct', 17], ['required', 16]]

TOP 10 Adjectives First Website

[['least', 58], ['able', 51], ['More', 34], ['local', 29], ['international', 23], ['foreign', 15], ['appropriate', 14], ['relevant', 14], ['professional', 13], ['advanced', 11]]
[['NUST', 35], ['Sciences', 17], ['Admissions', 10], ['ENGINEERING', 10], ['Pakistan', 9], ['SCHOOL', 9], ['National', 8], ['University', 8], ['Technology', 8], ['Achievements', 8]]
[['offers', 3], ['visits', 2], ['retains', 2], ['stands', 2], ['are', 2], ['selected', 1], ['awarded', 1], ['Summer', 1], ['transferred', 1], ['Closed', 1]]
[['more', 3], ['top', 2], ['More', 2], ['economic', 2], ['recent', 1], ['many', 1], ['dynamic', 1], ['higher', 1], ['various', 1], ['robust', 1]]
```

Top 10 nouns, adjectives, and verbs with their weights of second university website (NUST).

```
TOP 10 Nouns of second Website

[['NUST', 35], ['Sciences', 17], ['Admissions', 10], ['ENGINEERING', 10], ['Pakistan', 9], ['SCHOOL', 9], ['National', 8], ['University', 8], ['Technology', 8], ['Achievements', 8]]

TOP 10 Verbs Second Website

[['offers', 3], ['visits', 2], ['retains', 2], ['stands', 2], ['are', 2], ['selected', 1], ['awarded', 1], ['Summer', 1], ['transferred', 1], ['Closed', 1]]

TOP 10 Adjectives Second Website

[['more', 3], ['top', 2], ['More', 2], ['economic', 2], ['recent', 1], ['many', 1], ['dynamic', 1], ['higher', 1], ['various', 1], ['robust', 1]]
[['LUMS', 35], ['School', 19], ['Lorem', 18], ['dolor', 18], ['ipsum', 17], ['sit', 16], ['amet', 16], ['consectetur', 16], ['adipisc', 15], ['Read', 13]]
[['Submitted', 10], ['are', 4], ['have', 3], ['get', 3], ['Give', 2], ['create', 2], ['committed', 2], ['explore', 2], ['achieve', 2], ['Imagine', 1]]
[['more', 7], ['More', 4], ['Last', 3], ['extraordinary', 2], ['liberal', 2], ['global', 2], ['main', 1], ['future', 1], ['adipisc', 1], ['next', 1]]
```

Top 10 nouns, adjectives, and verbs with their weights of third university website (LUMS).

```
TOP 10 Nouns of third Website

[['LUMS', 35], ['School', 19], ['Lorem', 18], ['dolor', 18], ['ipsum', 17], ['sit', 16], ['amet', 16], ['consectetur', 16],
['adipisc', 15], ['Read', 13]]

TOP 10 Verbs third  Website

[['Submitted', 10], ['are', 4], ['have', 3], ['get', 3], ['Give', 2], ['create', 2], ['committed', 2], ['explore', 2], ['ach
ieve', 2], ['Imagine', 1]]

TOP 10 Adjectives third  Website

[['more', 7], ['More', 4], ['Last', 3], ['extraordinary', 2], ['liberal', 2], ['global', 2], ['main', 1], ['future', 1], ['a
dipisc', 1], ['next', 1]]
```
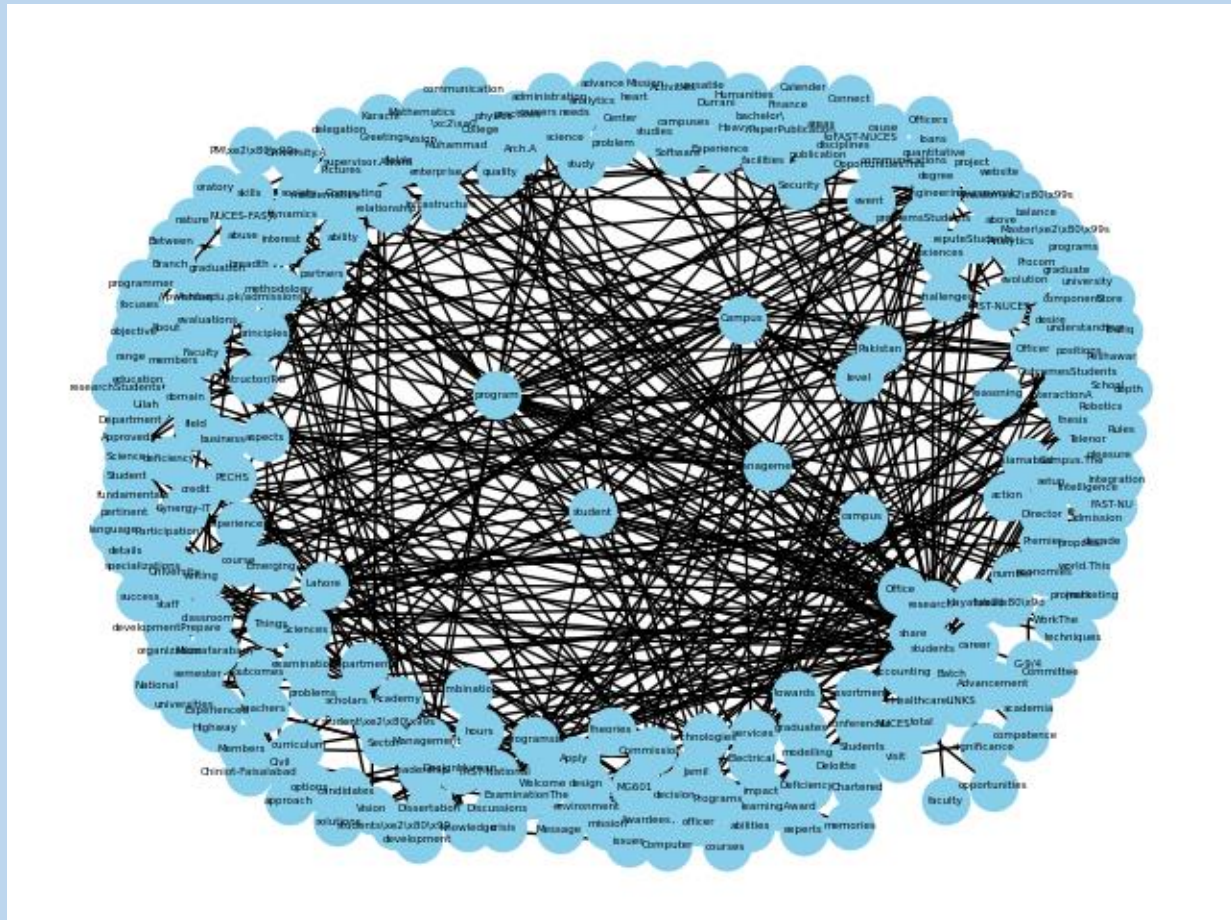
SECTION 03:

1. **GRAPH**
2. **CONNECTED COMPONENTS**
3. **10 NOUNS WITH HIGHEST DEGREE**
4. **QUALITY NOUN NODE**

Section 03:

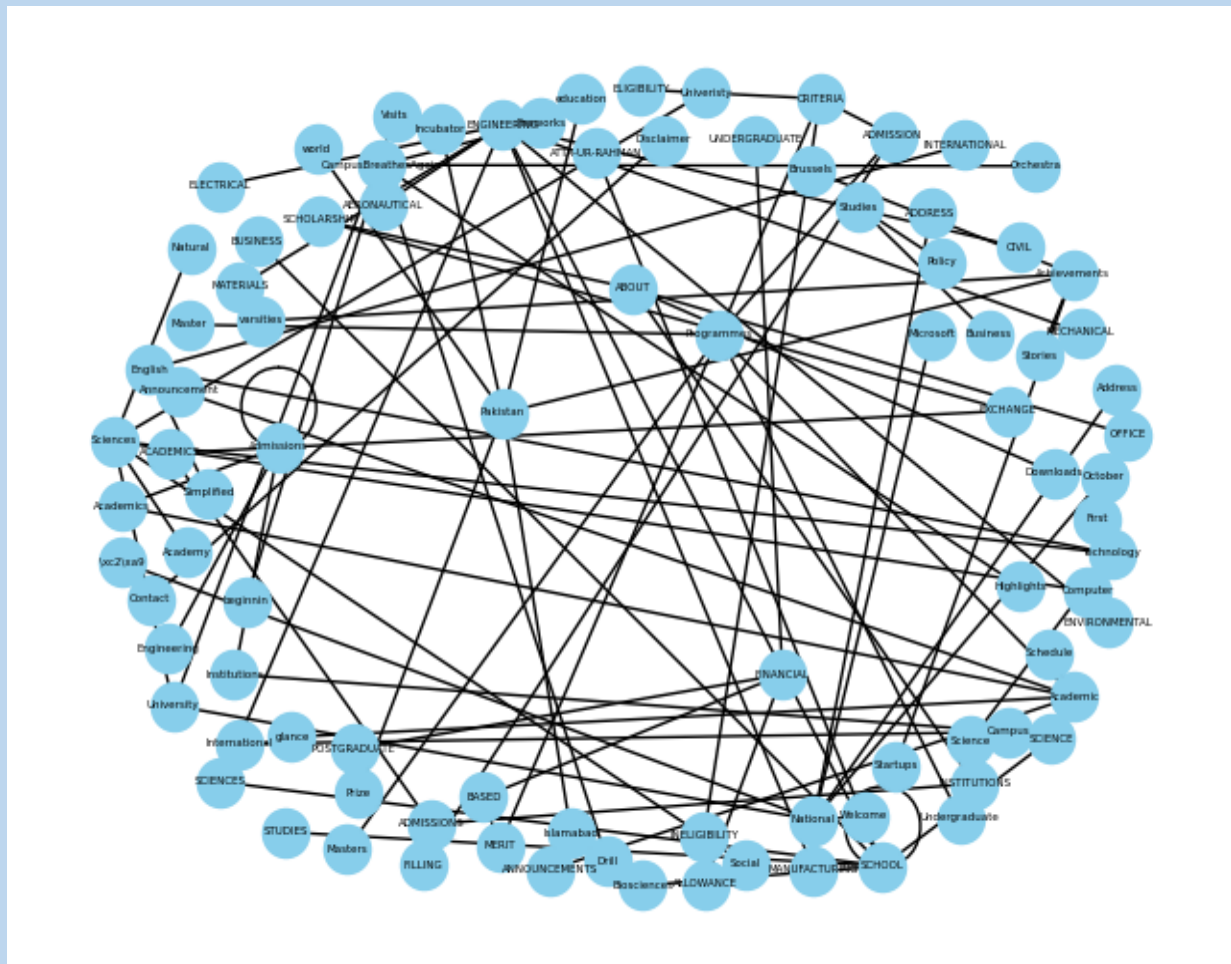### 1. Connected graph

NUCES Website:

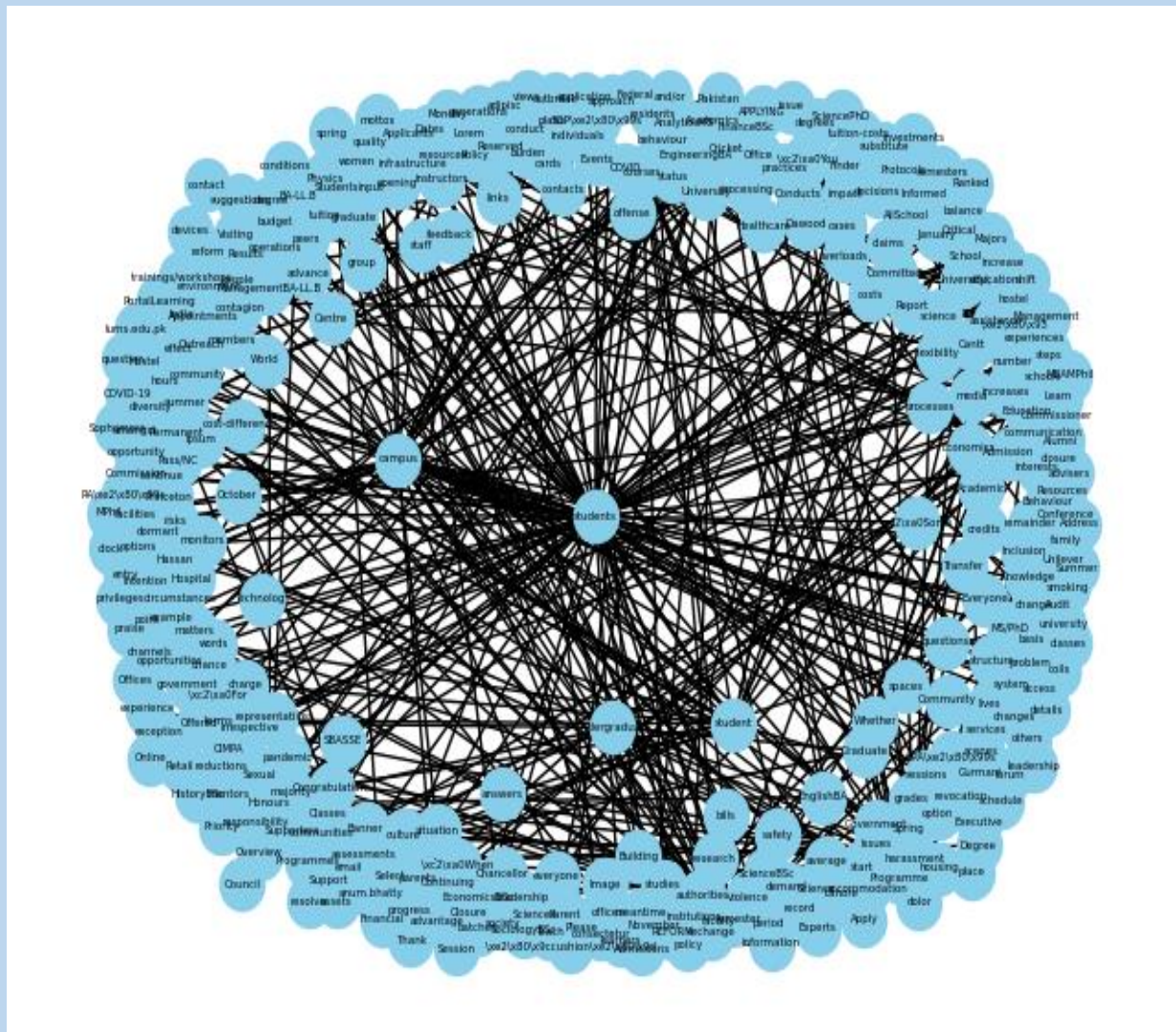First image shows the connected component those are the top 10 occurred nouns in the first website.

The second picture is the pictorial representation of the nouns collected from the NUCES website.

WEBSITE 02 (NUST):

This the noun graph of the second website.

Noun connected graph of the third website (LUMS).

2. 10 nouns with the highest weight

Top 10 nouns of NUCES website (1ST)

```
Number of Connected Componets is:1
########### TOP APPEARING IN THE GRAPH #########
('research', 42)
('students', 36)
('program', 34)
('management', 27)
('courses', 24)
('skills', 23)
('knowledge', 22)
('Campus', 21)
('Peshawar', 18)
('degree', 18)
```

Top 10 nouns of NUST website (2nd)

```
Number of Connected Componets is:8
########### TOP APPEARING IN THE GRAPH #########
('SCHOOL', 9)
('ENGINEERING', 9)
('National', 7)
('Sciences', 7)
('Pakistan', 6)
('Academic', 6)
('Admissions', 5)
('Achievements', 5)
('Programmes', 5)
('CampusBreathesAgain', 5)
```

Top 10 nouns of LUMS website (3RD)

```
Number of Connected Componets is:3
########### TOP APPEARING IN THE GRAPH #########
('students', 87)
('campus', 53)
('harassment', 30)
('courses', 28)
('student', 23)
('semester', 22)
('faculty', 19)
('Pakistan', 16)
('Admissions', 13)
('University', 13)
```

3. "Quality" noun
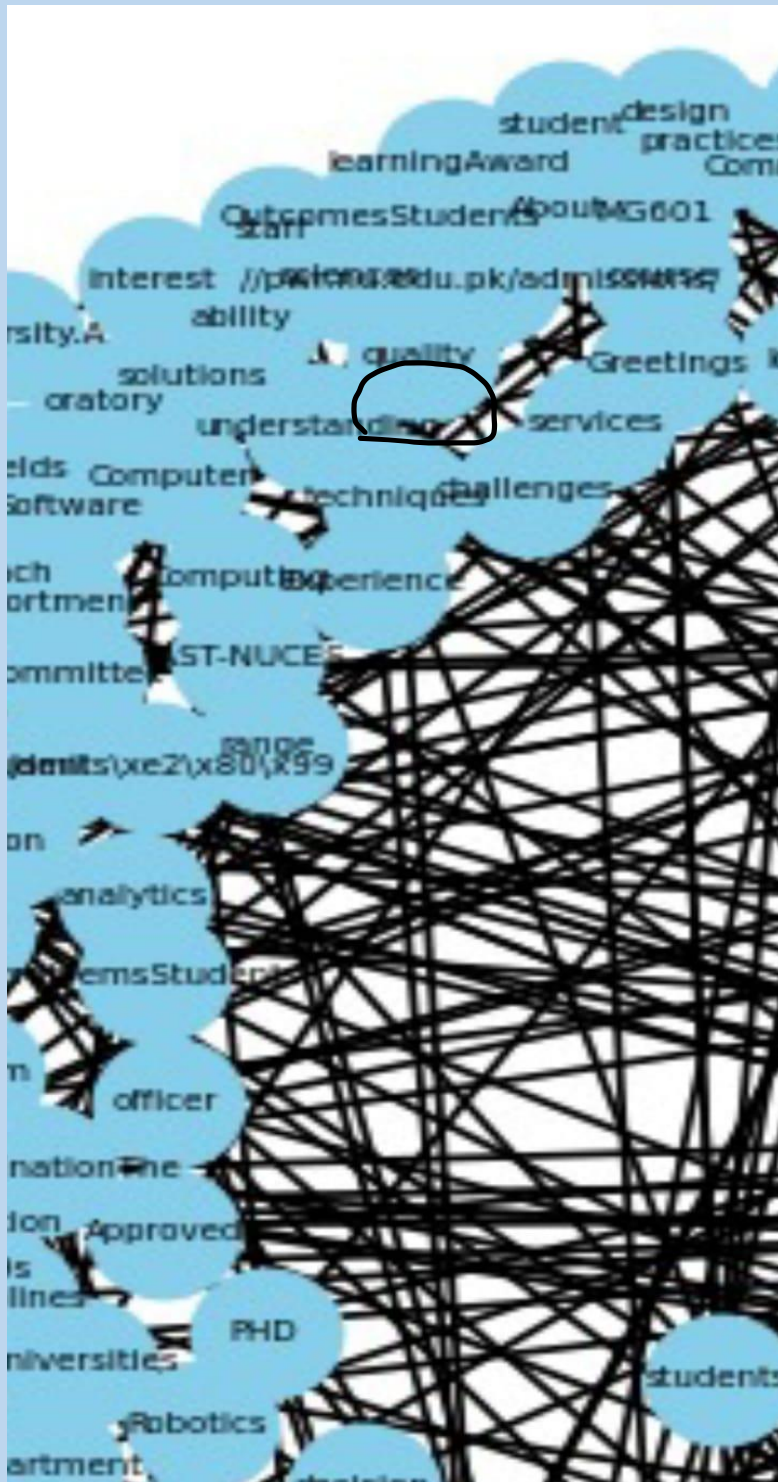
Quality noun was available in all the webpages of given websites.

But if we take NUCES website and checked the connectivity of "quality" with different nouns.

Quality was connected to

1. Services
2. Residence
3. Loans
4. Educations
5. Facilities
6. Security
7. Thesis
8. Study