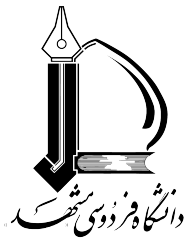




# **Data Analysis Using Image Processing, Text Processing, and Social Network Analysis**

**Name: Amirali Satvati**

**Instructor: Mohammad Arashi**



January 5, 2025

## Contents

<b>1</b>	<b>Introduction: Data Analysis Using Image Processing, Text Processing, and Social Network Analysis</b>	<b>2</b>
1.1	Image Editing and Animation Creation . . . . .	2
1.2	Text Analysis and Word Cloud Generation . . . . .	2
1.3	Social Network Analysis . . . . .	2
<b>2</b>	<b>Image Animation Processor</b>	<b>2</b>
<b>3</b>	<b>Word Cloud Builder</b>	<b>4</b>
<b>4</b>	<b>Network Graph Analyzer</b>	<b>6</b>
<b>5</b>	<b>References and Links</b>	<b>7</b>

# **1 Introduction: Data Analysis Using Image Processing, Text Processing, and Social Network Analysis**

In today's world, data analysis plays a vital role in extracting valuable insights from various data sources. These data can include images, texts, and social networks. This document introduces three different methods for data analysis, categorized as follows:

## **1.1 Image Editing and Animation Creation**

In the first part, using the powerful `magick` package in R, images are edited and converted into an animation in GIF format. This process includes resizing images, rotating, adding borders, and custom text. Finally, the images are transformed into a dynamic animation, which has numerous applications in scientific presentations and marketing.

## **1.2 Text Analysis and Word Cloud Generation**

In the second part, Persian poems are scraped from the `Ganjoor` website using web scraping and text processing. After removing stop words and symbols, a word cloud is generated. This word cloud uses various colors and shapes to visually represent frequent words and can be highly effective in text analysis and engaging audiences.

## **1.3 Social Network Analysis**

In the third part, social interaction data are read from a text file, and a social network graph is created using the `igraph` package in R. This graph models user interactions, examining various node and edge properties such as node degree. Finally, the graph is visualized with attractive graphical settings to make network information analyzable visually.

# **2 Image Animation Processor**

In this section, images are processed and converted into an animation.



Figure 1: Input image for processing

## R Code

```

1 library(magick)
2 image2 <- image_read("C:/Users/WIN10/Desktop/2.jpg")
3 image3 <- image_read("C:/Users/WIN10/Desktop/3.jpg")
4 image <- image_read("C:/Users/WIN10/Desktop/1.jpg")
5
6 image <- image_scale(image, "x500")
7 image <- image_rotate(image, degrees = 0)
8 image <- image_border(image_background(image, "hotpink"), "#000000", "0
  x5")
9 image <- image_annotate(image, "The Third Data Science Seminar", size =
  29, gravity = "east", color = "blue")
10
11 img <- c(image, image2, image3)

```

```

12 image_animate(image_scale(img, "200x200"), fps = 1, dispose = "
    background")
13
14 image_write(image_animate(image_scale(img, "200x200"), fps = 1, dispose
    = "background"), "C:/Users/WIN10/Desktop/imageanalysis.gif")

```

Listing 1: Image Animation Processor Code



Figure 2: Processed image with R

### 3 Word Cloud Builder

This section focuses on text analysis and word cloud generation.

#### R Code for Word Cloud Generation

```

1 library(tm)
2 library(SnowballC)
3 library(wordcloud)
4 library(wordcloud2)
5 library(RColorBrewer)
6 library(rvest)
7
8 url <- "https://ganjoor.net/hafez/ghazal"
9 page <- read_html(url)
10 text <- page %>% html_nodes("div.poem") %>% html_text()
11
12 docs <- Corpus(VectorSource(text))
13
14 toSpace <- content_transformer(function(x, pattern) gsub(pattern, " ",
15   x))
16 docs <- tm_map(docs, toSpace, "/")
17 docs <- tm_map(docs, toSpace, "@")
18 docs <- tm_map(docs, toSpace, "\\|")
19 docs <- tm_map(docs, removeNumbers)
20 docs <- tm_map(docs, content_transformer(tolower))
21 docs <- tm_map(docs, removePunctuation)
22 docs <- tm_map(docs, stripWhitespace)
23
24 #Here, the Persian font does not render properly because it cannot
25   read it. In the original Markdown code, this is formatted correctly
26   .
27
28 my_stopwords <- c(" ", " ", " ", " ", " ", " ", " ", " ", " ", " ", " ", " ", " ")
29 docs <- tm_map(docs, removeWords, my_stopwords)
30
31 dtm <- TermDocumentMatrix(docs)
32 m <- as.matrix(dtm)
33 v <- sort(rowSums(m), decreasing = TRUE)
34 d <- data.frame(word = names(v), freq = v)
35
36 wordcloud2(data = d, size = 1.5, color = "random-light",
37   backgroundColor = "black", shape = "diamond")

```

Listing 2: Word Cloud Builder Code



Figure 3: Word Cloud

## 4 Network Graph Analyzer

In this section, social network analysis and the related graph are presented.

```

1 library(igraph)
2 data <- read.table("C:\\Users\\WIN10\\Desktop\\SocialNetworkAnalysis.
      txt", sep = ",", header = TRUE)
3 y <- data.frame(data$first, data$second)
4
5 net <- graph.data.frame(y, directed = TRUE)
6 V(net)$label <- V(net)$name
7 V(net)$degree <- degree(net)
8
9 layout <- layout_with_fr(net) * 0.3
10 plot(net,
11       vertex.color = rainbow(52, alpha = 0.8), vertex.size = 20,
12       vertex.label.color = "black", edge.color = "gray80",
13       layout = layout, main = "Social Network Analysis")

```

Listing 3: Network Graph Analyzer Code

### Social Network Analysis

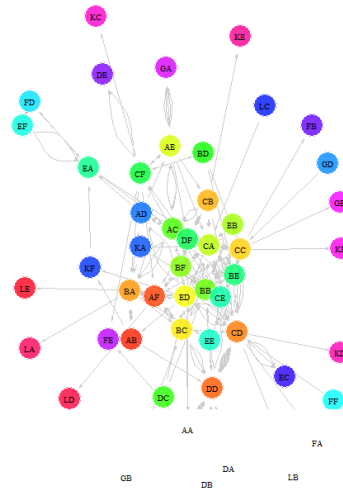


Figure 4: Social Network Analysis

## 5 References and Links

I deeply appreciate you accompanying me on this journey so far.

For more information, you can visit the following links: my [LinkedIn](#) or [GitHub](#) page. I warmly invite you to explore them for accessing files or connecting further.

- [GitHub](#)
- [LinkedIn](#)