Closer to Christ

Sam Lee

A Stylometric Analysis of the Four Gospels: The Names of Christ and Other Distinguishing Features

In this analysis I will attempt to perform a rudimentary stylometric analysis on the four gospels of the King James Version of the Bible. In this document I will examine which names of Christ show up most often in the gospels and whether there is a correlation between the style's of the different authors and the names of Christ along with other variables. We will create a multivariate regression model using R to evaluate my research questions.

Research Questions

- 1. Considering the four different authors of the four gospels used in the KJV version of the Bible each with their different focuses and audiences, is there a significant causal effect with one author using a particular name of Christ more than the other authors?
- 2. Can we predict which author wrote a specific line of text going off of the names of Christ alone?
- 3. Based on how the different authors wrote about Christ, what can we learn about Him and his nature?

Gathering the Data

To begin, I first created a large database in R of the KJV text of the four gospels. To do this, I web scraped the data directly off lds.org. I then formatted the data appropriately and cleaned all the text by removing all the footnotes.

```
library(tidyverse)
library(rvest)
library(stringr)
```

```
codes = c(
  "matt" = "Matthew",
  "mark" = "Mark",
  "luke" = "Luke",
  "john" = "John"
BookDBNames = c("Book", "Chapter", "VerseCount", "ChapterText")
dbName = "FourGospels.csv"
getChapterLength <- function(ldscode){</pre>
  chapters = NULL
  booklink = str_c("https://www.churchofjesuschrist.org/study/scriptures/nt/",
                   ldscode ,"?lang=eng")
  chapters = booklink %>% read_html() %>%
    html_nodes(".title") %>% length()
  #Close the unused the link
  url = url(booklink, "rb")
  close(url)
  print(str_c("Chapters in ", codes[ldscode] |> as.vector(), ": ", chapters))
  chapters
}
createDB <- function(ldscode=names(codes[1])){</pre>
  chapters = getChapterLength(ldscode)
  Verses = NULL
  text = c()
  verse_count = c()
  #Loop through each chapter found on lds.org and scrape the contents
  for(chapter in 1:chapters){
    link = str_c("https://www.churchofjesuschrist.org/study/scriptures/nt/",
                 ldscode ,"/",
                 chapter ,"?lang=eng")
    ChapterContents <- link %>% read_html() %>%
      html_nodes(".verse") %>%
```

```
html_text()
#print(ChapterContents)
#Get number of verses
verses <- link %>% read html() %>%
 html nodes(".verse-number") %>%
 html_text()
verse_count = c(verse_count, length(verses))
print(str_c("Verses in ",
            codes[ldscode] |> as.vector(), " ",
            chapter, ": ", length(verses)))
#Truncate the verse numbers at the beginning
ChapterContents = str_replace(ChapterContents, verses, "") |>
 str_replace_all(fixed("¶ "), "")
#Get rid of the footnotes (clean the data)
footnotes <- link %>% read html() %>%
 html_nodes(".study-note-ref") %>%
 html_text()
#prime the data for the loop
#get rid of the footnotes attached to parentheses
ChapterContents = gsub(".\\(", "(", ChapterContents)
#get the words with footnotes that haven't been cleaned yet
footnotes = footnotes[!grepl("\\(|\\)", footnotes)]
#An algorithm to find the words that have footnote letters attached to them
for(row in 1:length(ChapterContents)){
 fns = unlist(str_match_all(ChapterContents[row], footnotes))
 for(word in fns){
   #Remove the footnote for every word that has a footnote attached
   ChapterContents[row] = str_replace(
      ChapterContents[row], word, substr(word, 2, nchar(word))
   )
 }
}
```

```
#Cleaned text
    text = c(text, paste(ChapterContents, collapse = " "))
    vtable = tibble(Book = codes[ldscode],
                    Chapter = chapter,
                    Verse = 1:length(verses),
                    VerseText = ChapterContents)
    if(is.null(Verses)){
      Verses = vtable
    } else {
      Verses = Verses %>% rbind(vtable)
    #Close the unused the link after we're done
    url = url(link, "rb")
    close(url)
  }
  #Return a data base with the respective data vectors
  db = tibble(
    Book = codes[ldscode],
    Chapter = 1:chapters,
    VerseCount = verse_count,
    ChapterText = text
  ) %>% setNames(BookDBNames)
 return(db %>% left_join(Verses))
}
if (file.exists(dbName)) {
 FourGospels <- read_csv(dbName)</pre>
} else {
  FourGospels <- map_df(names(codes), createDB)</pre>
  FourGospels %>%
    mutate(ChptCharacterCount = nchar(ChapterText),
           ChptWordCount = str_count(ChapterText, '\\w+'),
           ChptAvgWordLength = nchar(
             str_replace_all(ChapterText, "[[:punct:]]", ""))/ChptWordCount,
           VerseCharacterCount = nchar(VerseText),
           VerseWordCount = str_count(VerseText, '\\w+'),
           VerseAvgWordLength = nchar(
```

```
str_replace_all(VerseText, "[[:punct:]]", "")
                                    )/VerseWordCount) %>%
            group_by(Book) %>% mutate(
                  PctOfBook = VerseCharacterCount/sum(VerseCharacterCount),
                  AuthorsAvgWordLength = sum(PctOfBook*VerseAvgWordLength)) %>%
            group_by(Book, Chapter) %>%
                  mutate(
                        ChptPctOfBook = sum(PctOfBook),
                        UniqueWordsinChp = str_replace_all(ChapterText, "[[:punct:]]", "") %>%
                               toupper() %>%
                              str_split(" ") %>%
                              unlist() %>%
                              unique() %>%
                             length()
                  ) -> FourGospels
      FourGospels %>% group_by(Book, Chapter, Verse) %>%
            mutate(
                  UniqueWordsinVerse = str_replace_all(VerseText, "[[:punct:]]", "") %>%
                        toupper() %>%
                        str_split(" ") %>%
                        unlist() %>%
                        unique() %>%
                        length()
            ) -> FourGospels
      FourGospels %>% group_by(Book) %>%
            mutate(
                  UniqueWordsinBook = str replace_all(ChapterText, "[[:punct:]]", "") %%
                        toupper() %>%
                        str_split(" ") %>%
                        unlist() %>%
                        unique() %>%
                        length()
            ) -> FourGospels
      #should sum to 1
      FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(FourGospels(F
      write_csv(FourGospels, dbName)
}
```

Here is the resulting database

Book	Chapter	VerseCount	Verse	VerseText
Matthew	1	25	1	The book of the generation of Jesus
				Christ, the son of David, the son of
				Abraham.
Matthew	1	25	2	Abraham begat Isaac; and Isaac begat
				Jacob; and Jacob begat Judas and his
				brethren;
Matthew	1	25	3	And Judas begat Phares and Zara of
				Thamar; and Phares begat Esrom; and
				Esrom begat Aram;
Matthew	1	25	4	And Aram begat Aminadab; and
				Aminadab begat Naasson; and Naasson
				begat Salmon;
Matthew	1	25	5	And Salmon begat Booz of Rachab; and
				Booz begat Obed of Ruth; and Obed
				begat Jesse;

...with 3,774 more rows (other statistical variables omitted for the purposes of display).

Included in the table above are also some basic statistics. I collected data on 3779 verses, with 1071 verses in Matthew, 678 verses in Mark, 1151 verses in Luke, and 879 verses in John.

There are 23727 words in Matthew, 15186 words in Mark, 25986 words in Luke, and 19125 words in John.

Here's how average word length per author breaks down. While all are really close, the Book of Matthew has, on average, the longest words at 5.054 letters per word, followed by Mark (5.021), Luke (4.993), and then John (4.928).

The Book of Luke contains the most unique words at 2401 words, followed by Matthew (2130), Mark (1681), and then John (1403).

The Names of Christ

The most important part of this analysis is to actually find out how Christ is referenced by the four gospel authors. To do this, I compiled a list of 64 names and titles that Christ could possibly be called. I consulted these two sources from the Church website: The Names of Christ, 50 Names for Jesus.

Not all of these names were found in the four gospels of course so when I narrowed down my search and found the names that were actually referenced by at least one of the four gospel authors as Jesus Christ, I found 26 names of Christ.

I created columns for these in the database and created a short script that found the frequency at which they appeared in the text.

```
#Get the names of Christ
names_url1 = str_c("https://site.churchofjesuschrist.org/",
  "study/ftsoy/2022/12/03-the-names-of-christ?lang=eng")
names1 = names_url1 %>% read_html() %>%
  html nodes("h2") %>% html text()
names1 = unlist(strsplit(names1, ",\\s*"))
names_url2 = str_c("https://www.churchofjesuschrist.org/comeuntochrist/",
  "believe/jesus/50-names-for-jesus")
names2 = names_url2 %>% read_html() %>%
 html_nodes("span.typography_h4") %>%
 html text()
names2 = gsub("^\d+[--.]*\s*|(?<=\S)\\s*\\d+[--.]*\\s*|\\s*", "",
              names2, perl = TRUE)
names2 = unlist(strsplit(names2, ",\\s*"))
#Combine the names and add in alternative spellings for the KJV
names = unique(c(names1, names2, "Saviour", "Jesus", "Jesus Christ"))
#Remove these names because Christ is not being referred to here in the gospels
#when these are mentioned is mentioned
#Additionally Good Shepherd is a duplicate
names = names[!(names %in% c("Rock", "Friend", "Judge", "Good Shepherd"))]
#Find the number of occurrences for each word through the gospels
#and assign it its own variable name in FourGospels
for(i in 1:length(names)){
  name = names[i]
  if(name %in% c("I Am", "Wonderful", "Counsellor", "Master",
                 "High Priest", "Prophet", "Judge", "Branch",
                 "The Way", "Lord", "The Word")){
    #In this case, case sensitivty matters
    counts = str_count(FourGospels$VerseText, name)
  } else {
    #Find the occurrences of Jesus Christ in the gospels, ignoring case
    counts = str_count(toupper(FourGospels$VerseText), toupper(name))
  }
  FourGospels[name] = counts
```

Here are the 26 Names of Christ Found in the Gospels

Index	Name	Frequency
1	The Good Shepherd	3
2	Light of the World	3
3	God of Israel	2
4	Emmanuel	1
5	Lamb of God	2
6	Bread of Life	2
7	Lord	211
8	Son of the Living God	2
9	Only Begotten Son	3
10	Beloved Son	7
11	Master	55
12	King of the Jews	18
13	Prophet	1
14	Teacher	1
15	Messias	2
16	True Vine	1
17	Bridegroom	15
18	Dayspring	1
19	Son of Man	85
20	Carpenter	2
21	King of Israel	4
22	Christ	62
23	Saviour	3
24	Jesus	625
25	Jesus Christ	5

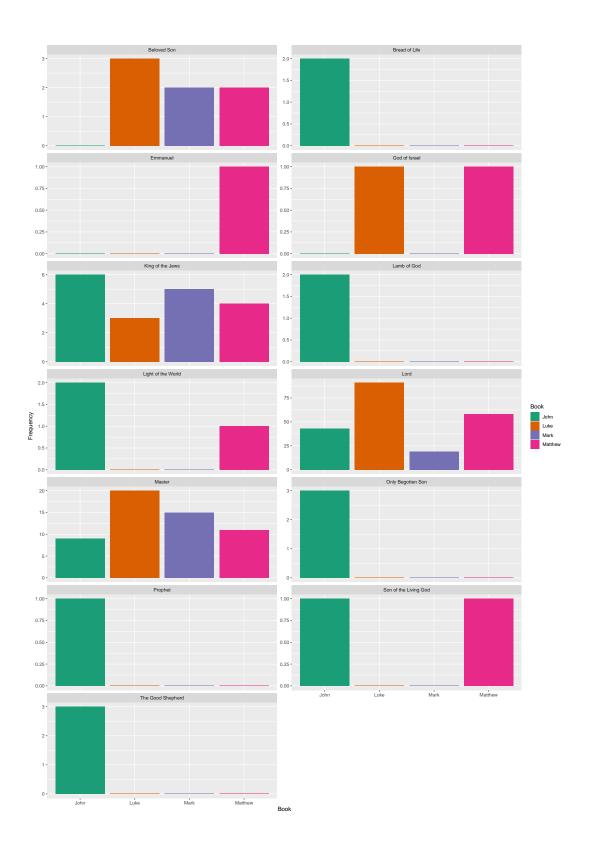
Analysis

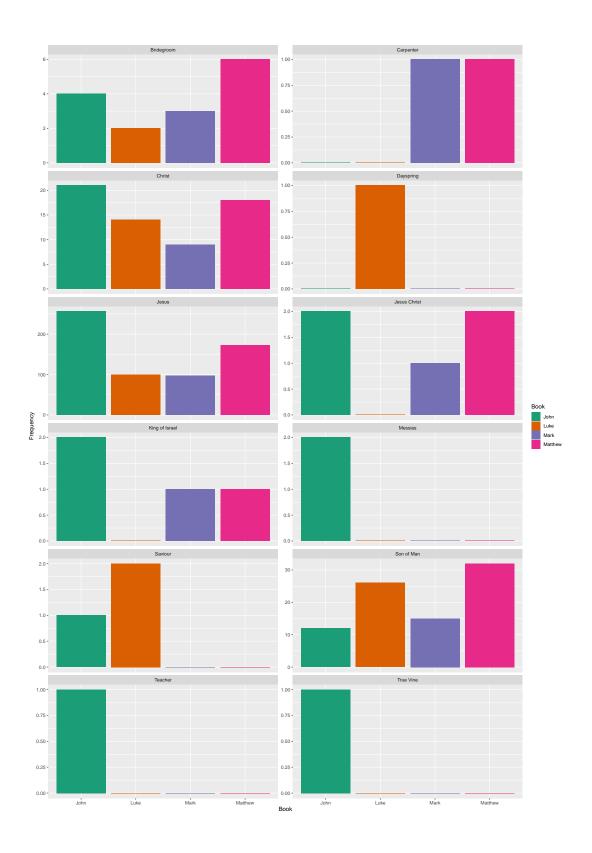
Now that we have collected the data and broken down the names of Christ found within the gospels, we can analyze where exactly they are found in the gospels and how they are used.

Here is a net summary of where each of these names fall in regard to authorship:

```
#Analysis of the names
word_columns <- FourGospels[, names]
#Calculate the sum of each column</pre>
```

```
word_sums <- colSums(word_columns)</pre>
#Keep the columns if there is at least 1 occurrence
filtered_words <- word_columns[, word_sums >= 1]
names_authorship = FourGospels %>% group_by(Book) %>%
    across(names(filtered_words), ~ sum(.))
  )
num_words = length(colnames(filtered_words))
half = ((num_words/2) %>% ceiling())
name_distribution1 = FourGospels %>% group_by(Book) %>%
  summarize(
    across(names(filtered_words)[1:half], ~ sum(.))
  ) %>% pivot_longer(cols = colnames(filtered_words)[1:half],
                     values_to = "Frequency")
name_distribution2 = FourGospels %>% group_by(Book) %>%
  summarize(
    across(names(filtered_words)[(half+1):num_words], ~ sum(.))
  ) %>% pivot_longer(cols = colnames(filtered_words)[(half+1):num_words],
                     values_to = "Frequency")
```





Random Forest Model

To answer one of my questions, I created a random forest model to predict the authorship of any given verse given the names of Christ alone.

```
library(randomForest)

# Prepare the data - Random Forest Model
word_columns <- FourGospels[, colnames(filtered_words)]

books <- as.factor(FourGospels$Book)
data <- cbind(word_columns, books)

sample_p = 0.9</pre>
```

I randomly selected 90% of the verses from the gospels for the model to be trained on, and left the remainder of verses to be left for trial.

```
train_indices <- sample(1:nrow(data), sample_p * nrow(data))
train_data <- data[train_indices, ]
test_data <- data[-train_indices, ]

colnames(train_data) <- gsub(" ", "_", colnames(train_data))
colnames(test_data) <- gsub(" ", "_", colnames(test_data))

# Build the Random Forest model
rf_model <- randomForest(books ~ ., data = train_data)</pre>
```

To see how well the unique names of Christ used supposedly uniquely by the four different authors, we will see if it can predict the authorship of the 10% remaining verses.

Here's the model against the predictions:

```
# Model Evaluation
predictions <- predict(rf_model, test_data)
library(caret)
cm = confusionMatrix(predictions, test_data$books)</pre>
```

	John	Luke	Mark	Matthew
John	20	11	6	20
Luke	55	120	59	87
Mark	0	0	0	0
Matthew	0	0	0	0

The numbers along the diagonal are the accurate predictions, that means a 37.037037% accuracy. It's reasonable to conclude that this model is not very good at predicting the authorship on the basis of the names of Christ alone. What is interesting is that this model typically predicts very few (if any) verses in Mark. Mark may not be "unique" enough for the model to distinguish it from the other authors. This refortifies the theory that Mark was a common source for both Luke and Matthew.

Adjusting the Model: Unique Features to the Four Gospels

Here in this next bit, I am going to analyze a couple more frequencies of words that appear throughout the gospels and see if this can me understand the distinguished authorship.

The words I chose to analyze are love, pray, repentance, and faith (and their associated conjugates). These words show up pretty universally in the gospels and I wanted to see if there was a connection between authorship, how these authors used these words, and how they mentioned Christ.

I first added these words to my database:

```
#Love
#includes love, loves, loved, etc. because they all start w/ love
love_counts = str_count(toupper(FourGospels$VerseText), toupper("love")) +
    str_count(toupper(FourGospels$VerseText), toupper("loving"))
FourGospels["Love"] = love_counts

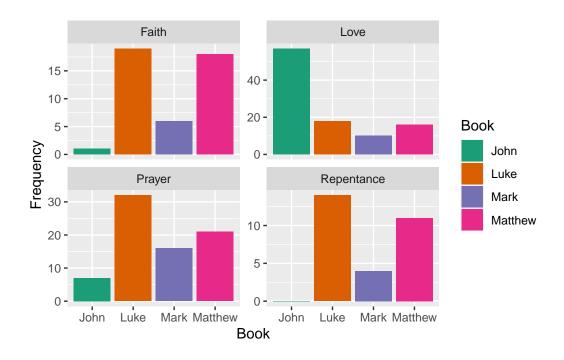
#Repent, Repentance
#includes repent, repents, repentance, repented, etc. because they all start w/ repent
repentance_counts = str_count(toupper(FourGospels$VerseText), toupper("repent"))
FourGospels["Repentance"] = repentance_counts

#Faith
faith_counts = str_count(toupper(FourGospels$VerseText), toupper("faith"))
FourGospels["Faith"] = faith_counts

#Pray, prayer
#includes pray, prays, prayed, prayer, etc. because they all start w/ pray
prayer_counts = str_count(toupper(FourGospels$VerseText), toupper("pray"))
FourGospels["Prayer"] = prayer_counts
```

Next, similar to what we did above, let's see how these words are distributed throughout the Gospel:

```
special_words = c("Love", "Repentance", "Faith", "Prayer")
         word_distribution = FourGospels %>% group_by(Book) %>%
               summarize(
                      across(special_words, ~ sum(.))
               ) %>% pivot_longer(cols = special_words, values_to = "Frequency")
Warning: There was 1 warning in `summarize()`.
i In argument: `across(special_words, ~sum(.))`.
Caused by warning:
! Using an external vector in selections was deprecated in tidyselect 1.1.0.
i Please use `all_of()` or `any_of()` instead.
       data %>% select(special_words)
       # Now:
       data %>% select(all_of(special_words))
See <a href="feether: 10%">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/faq-external-vector.html>">https://tidyselect.r-lib.org/reference/
        word_distribution %>% ggplot(mapping=aes(x=Book, y=Frequency, fill=Book)) +
               geom_bar(stat = "identity") +
               facet_wrap(name ~ ., ncol=2, scales = "free_y") +
               scale_fill_brewer(type = "qual", palette = 2)
```



At first glance, the distributions for faith and prayer look really similar; The distributions for love and repentance look significant.

I performed a simple analysis of variance test to test the difference of means across these word distributions with Book as a factor

```
word.lm.Love = lm(Love ~ Book, data=FourGospels)
word.lm.Repentance = lm(Repentance ~ Book, data=FourGospels)
word.lm.Prayer = lm(Prayer ~ Book, data=FourGospels)
word.lm.Faith = lm(Faith ~ Book, data=FourGospels)
anova(word.lm.Love)
```

Analysis of Variance Table

```
Response: Love

Df Sum Sq Mean Sq F value Pr(>F)

Book 3 1.665 0.55496 12.572 0.00000003547 ***

Residuals 3775 166.636 0.04414

---

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

anova(word.lm.Repentance) Analysis of Variance Table Response: Repentance Df Sum Sq Mean Sq F value Pr(>F) Book 3 0.0843 0.0281061 3.4568 0.01576 * Residuals 3775 30.6931 0.0081306 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 anova(word.lm.Prayer) Analysis of Variance Table Response: Prayer Df Sum Sq Mean Sq F value Pr(>F) 3 0.206 0.068768 3.0808 0.02636 * Residuals 3775 84.265 0.022322 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 anova(word.lm.Faith) Analysis of Variance Table Response: Faith Df Sum Sq Mean Sq F value Pr(>F) 3 0.158 0.052697 4.0327 0.007113 ** Residuals 3775 49.330 0.013067

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

All distributions are significantly uneven (all p-values < 0.05), which means there is significant evidence to suggest that different authors talked about these topics differently, most notably, there isn't a single occurrence of "repentance" or any of its conjugations in the Book of John.

To increase model accuracy, adjusting for other stylometric factors including average word length and number of unique words, will that be enough to adjust the model's predictive power significantly?

If we adjust the Random Forest model previously including these effects, will the model improve?

```
FourGospelsCondensed = FourGospels %>%
  filter(rowSums(across(colnames(filtered_words))) > 0)
word columns <- FourGospelsCondensed[, c("VerseAvgWordLength", "UniqueWordsinVerse",</pre>
                                    colnames(filtered words), special words)]
books <- as.factor(FourGospelsCondensed$Book)</pre>
data <- cbind(word_columns, books)</pre>
sample_p = 0.9
train_indices <- sample(1:nrow(data), sample_p * nrow(data))</pre>
train_data <- data[train_indices, ]</pre>
test_data <- data[-train_indices, ]</pre>
colnames(train_data) <- gsub(" ", "_", colnames(train_data))</pre>
colnames(test_data) <- gsub(" ", "_", colnames(test_data))</pre>
# Build the Random Forest model
rf_model <- randomForest(books ~ ., data = train_data)</pre>
predictions <- predict(rf_model, test_data)</pre>
library(caret)
cm2 = confusionMatrix(predictions, test_data$books)
```

To try this out, could you guess the authorship of this verse (randomly selected from the same data that the model will use)?

Here's a random verse from the predictive set:

And as they came down from the mountain, Jesus charged them, saying, Tell the vision to no man, until the Son of man be risen again from the dead.

Here's what the model predicts:

Matthew

Here's the actual verse: Matthew 17:9

Here is the new prediction output:

	John	Luke	Mark	Matthew
John	28	7	15	16
Luke	4	9	7	6
Mark	0	0	0	0
Matthew	0	2	2	1

Sometimes it's due to random chance and the data the model is trained on, but it looks like this model performs more accurately than the previous model with an accuracy of 39.1752577%. There's still a lot of improvement to do but, this shows that authorship cannot be defined so simply. It certainly not contained within a few words: This is especially the case when authors, such as Matthew and Luke build off another, such as Mark. The name of Christ is very profound and no single gospel author had a monopolistic use on a term or series of terms.

Limitations

While this is a simple stylometric analysis of the four gospels and how they use the names of Christ, this way of particularly studying the gospels from an analytical standpoint has several obvious limitations. The greatest obstacle with this kind of analysis is that I am working with a translation, albeit, a somewhat removed in some ways and by no means perfect translation. The gospel writers wrote the original transcripts in Greek, and I am analyzing these texts in the same KJV English translation. When we deal with translations, a lot of original meaning can be lost. Additionally, the same translation team of King James worked on translating the Bible, so even if the original authors of the gospels had distinct writing styles and that was apparent in the ways that they used the names of Christ among other things, it would be fair to reason that the same translation team of King James had similar translation styles throughout the process of translating the Bible. The distinguishing differences between the four authors have the potential to be lost in the tendency of the translators and their translation process. The other limitation with this study is the fact that a lot of the names of Christ might be hidden behind other first-person or third-person pronouns which I did not analyze. I analyzed the titles of Christ which while appeared throughout the text may not have yielded sufficient data. A more sophisticated approach would be to analyze the times when Jesus Christ is being referred to in the first or third person. This could yield more reliable data for a model to use.

Conclusion

With the model variables used: the names of Christ, a few unique words that are significantly used differently by the four different authors of the gospels, along with a couple of stylometric metric used for analysis, it proves to be very difficult to "find the signal in noise" to show empirically which author must have written a particular verse. We can show absolutely how and where these verses are found when we look at the names of Christ alone.

I found the verses in the Bible where the names of Christ appear at the highest frequency for the 26 titles I analyzed.

Here are the verses in Matthew:

Name	Verse	VerseText
Light of the World	Matthew 5:14	Ye are the light of the world. A city that is
		set on an hill cannot be hid.
God of Israel	Matthew 15:31	Insomuch that the multitude wondered,
		when they saw the dumb to speak, the
		maimed to be whole, the lame to walk, and
		the blind to see: and they glorified the God
		of Israel.
Emmanuel	Matthew 1:23	Behold, a virgin shall be with child, and
		shall bring forth a son, and they shall call
		his name Emmanuel, which being
		interpreted is, God with us.
Lord	Matthew 7:21	Not every one that saith unto me, Lord,
		Lord, shall enter into the kingdom of
		heaven; but he that doeth the will of my
		Father which is in heaven.
Son of the Living God	Matthew 16:16	And Simon Peter answered and said, Thou
		art the Christ, the Son of the living God.
Beloved Son	Matthew 3:17	And lo a voice from heaven, saying, This is
		my beloved Son, in whom I am well pleased.
Master	Matthew 8:19	And a certain scribe came, and said unto
		him, Master, I will follow thee whithersoever
0.03.5	3.5 3.4.20	thou goest.
Son of Man	Matthew 24:30	And then shall appear the sign of the Son of
		man in heaven: and then shall all the tribes
		of the earth mourn, and they shall see the
		Son of man coming in the clouds of heaven
<u> </u>	M 11 10 FF	with power and great glory.
Carpenter	Matthew 13:55	Is not this the carpenter's son? is not his
		mother called Mary? and his brethren,
IZ: CI 1	M 441 07 40	James, and Joses, and Simon, and Judas?
King of Israel	Matthew 27:42	He saved others; himself he cannot save. If
		he be the King of Israel, let him now come
		down from the cross, and we will believe
Jesus Christ	Matthew 1:1	him. The beak of the generation of Issue Christ
Jesus Christ	matthew 1:1	The book of the generation of Jesus Christ,
		the son of David, the son of Abraham.

There are no verses in Mark that contain a verse with the higest frequency of titles for Christ.

Here are the verses in Luke:

Name	Verse	VerseText
Dayspring	Luke 1:78	Through the tender mercy of our God;
		whereby the dayspring from on high hath
		visited us,
Christ	Luke 4:41	And devils also came out of many, crying
		out, and saying, Thou art Christ the Son of
		God. And he rebuking them suffered them
		not to speak: for they knew that he was
		Christ.
Saviour	Luke 1:47	And my spirit hath rejoiced in God my
		Saviour.

Here are the verses in John:

Name	Verse	VerseText
The Good Shepherd	John 10:11	I am the good shepherd: the good shepherd giveth his life for the sheep.
Lamb of God	John 1:29	The next day John seeth Jesus coming unto him, and saith, Behold the Lamb of God, which taketh away the sin of the world.
Bread of Life	John 6:35	And Jesus said unto them, I am the bread of life: he that cometh to me shall never hunger; and he that believeth on me shall never thirst.
Only Begotten Son	John 1:18	No man hath seen God at any time; the only begotten Son, which is in the bosom of the Father, he hath declared him.
King of the Jews	John 19:21	Then said the chief priests of the Jews to Pilate, Write not, The King of the Jews; but that he said, I am King of the Jews.
Prophet	John 7:40	Many of the people therefore, when they heard this saying, said, Of a truth this is the Prophet.
Teacher	John 3:2	The same came to Jesus by night, and said unto him, Rabbi, we know that thou art a teacher come from God: for no man can do these miracles that thou doest, except God be with him.
Messias	John 1:41	He first findeth his own brother Simon, and saith unto him, We have found the Messias, which is, being interpreted, the Christ.
True Vine	John 15:1	I am the true vine, and my Father is the husbandman.
Bridegroom	John 3:29	He that hath the bride is the bridegroom: but the friend of the bridegroom, which standeth and heareth him, rejoiceth greatly because of the bridegroom's voice: this my joy therefore is fulfilled.
Jesus	John 19:38	And after this Joseph of Arimathæa, being a disciple of Jesus, but secretly for fear of the Jews, besought Pilate that he might take away the body of Jesus: and Pilate gave him leave. He came therefore, and took the body of Jesus.

We can learn a lot about Christ in the verses about him. Each one of these verses reveals something about His divine nature. The verses in Matthew specifically refer to prophecies

being fulfilled when it mentions Old Testament names for Christ that the Israelites know him by such as the "God of Israel", "Emmanuel", and "King of Israel." The title, "Emmanuel" is one that uniquely appears in the Gospel of Matthew. I think this is the case because Matthew is appealing to a group of Israelites that is interested to seeing how this Jesus figure is fitting in with their modern perception of what the Messiah should be. Matthew shows here that Jesus fits all the traits of what Old Testament Jehovah was presumed to be.

While there are lots of references to Christ in the Gospel of Luke, one that is unique to Luke is the title of "Dayspring." I love Luke 1:78 which states, "Through the tender mercy of our God; whereby the dayspring from on high hath visited us." This really resonates with the intentions and attributes of Luke as a writer. He humanizes Jesus and assigns comforting traits to Him. Christ is one who gives us "tender mercy" and I love how Luke captures that with this unique title.

Finally, I love how John captures the essence of Jesus Christ with the unique titles he gives him. These titles don't necessarily appear in large frequencies as I showed throughout these four gospels, but upon a deeper analysis of these verses, there is a lot of richness contained within these verses in John. Titles like "The Good Shepherd" convey the meaning of Jesus's role as a guide in our lives, each one of gone astray. The title, "Bread of Life" also only appears in the gospel of John. John is extremely gifted in creating these perhaps single-use titles that stick forever because of their profound meaning. I think of this title every time I partake of the Sacrament. Following this theme, we see that other titles that are unique to the Gospel of John are "The True Vine," "Messias," "Teacher," "Prophet," the "Lamb of God," and the "Only Begotten Son." Again, none of these titles appear in large frequencies, but with each verse they deliver impactful messages about the divinity of Jesus Christ. Jesus Christ is all of those things on a profound and eternal level. He was the Lamb of God who sacrificed himself so that we might repent and return to our Heavenly Father. He is the Only Begotten Son of the Father. John making such a claim to his audience is incredibly profound. There is none other who is the "Only Begotten Son." There could be no other. Jesus Christ was the only one who could sacrifice himself for entire world for the pains, sufferings, inequalities, and sins of mankind. No other mortal on this earth could fulfill that kind of infinite burden.

When looking upon all these titles in a synoptic way in this context helps define Christ's eternal and divine nature. Christ is my Savior and Redeemer. The gospel authors testified of that. They testified to the comfort he brings to the broken-hearted and the strength he brings to the weak. They testified that he is the same Jehovah from the Old Testament. They testified that he is a real divine being who atoned for the sins of the world. They all testified of Christ.