

Jane Austen

Jingwen Xu

2020/11/20

```
tnum.authorize(ip="54.158.136.133")
```

```
## Available spaces: testspace, MEPED, alion-rf, shared-testspace, MSSP-1, wintergreen, testspa
ce, testspace, testspace, testspace, testspace, testspace, testspace, testspace, testspace, tes
tspace, testspace, testspace, testspace, testspace, testspace, testspace, testspace, testspace,
testspace, testspace, shared-testspace, hared-testspace
```

```
## Numberspace set to: shared-testspace
```

Sense and Sensibility

Create tags

The book *Sense and Sensibility* talks about marriage and love about a couple of sisters Elinor and Marianne, there are three main relationships running through the whole book: Elinor and Edward, Marianne and Willoughby, Marianne and Brandon. So I create tags with marriage and marry, and the names of three couples separately.

```
tnum.tagByQuery("*sense* has*= REGEXP(\"marriage|marry\")", adds = ("ref:marriage"))
```

```
## list(modifiedCount = 85, tagged = 85, removed = 0)
```

```
tnum.tagByQuery("*sense* has*= REGEXP(\"Elinor|Edward\")", adds = ("ref:EE"))
```

```
## list(modifiedCount = 851, tagged = 851, removed = 0)
```

```
tnum.tagByQuery("*sense* has*= REGEXP(\"Marianne|Willoughby\")", adds = ("ref:MW"))
```

```
## list(modifiedCount = 662, tagged = 662, removed = 0)
```

```
tnum.tagByQuery("*sense* has*= REGEXP(\"Marianne|Brandon\")", adds = ("ref:MB"))
```

```
## list(modifiedCount = 632, tagged = 632, removed = 0)
```

Make plots

Before making the plots, we create a function to turn all the chapters into a vector.

```

getchapter <- function(query_t2){
  n=length(query_t2)
  chapterlist <- 1:n
  for (i in 1:n) {
    chapterlist[i]<-
      as.character(substring(str_split(tnum.getAttrFromList(query_t2[i], "subject"), "[:/"])[[1]]
[[4], 9))
  }
  return(chapterlist)
}

```

Marriage and Marry

```
marry <- tnum.query("@ref:marriage",max = 120)
```

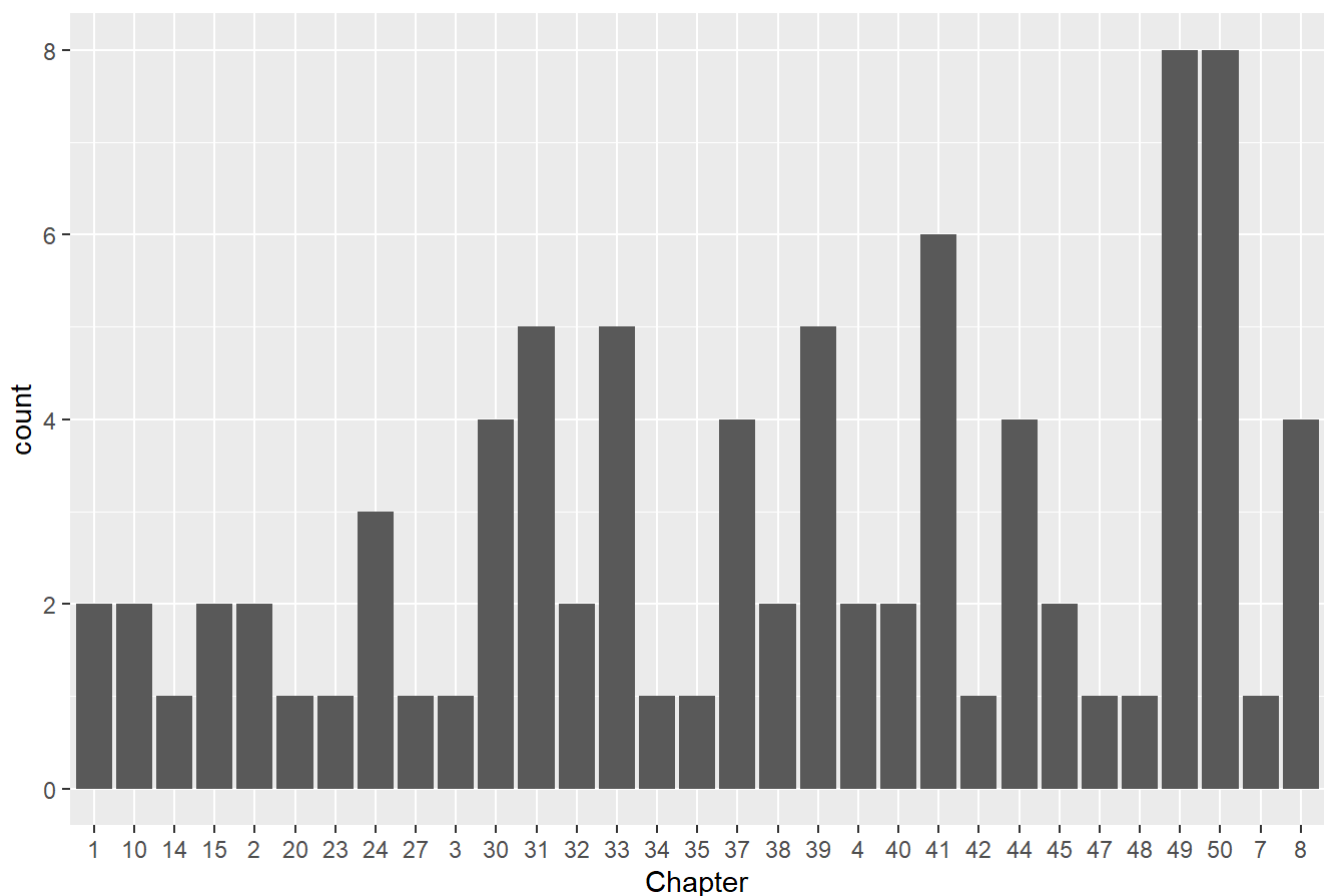
```
## Returned 1 thru 85 of 85 results
```

```

marrydf <- tnum.objectsToDf(marry)
ggplot()+geom_bar(mapping = aes(getchapter(marry)),stat = "count")+
  labs(x="Chapter",title = "Frequency of 'marry' and 'marriage'")

```

Frequency of 'marry' and 'marriage'



It's obvious that the frequency of marry and marriage in chapter 31-39 and chapter 49-50 is outstanding. And if we track the book *Sense and Sensibility*, we can see that the character's marriage crises mostly occur in chapter 31-39, and the couples get married in chapter 49-50

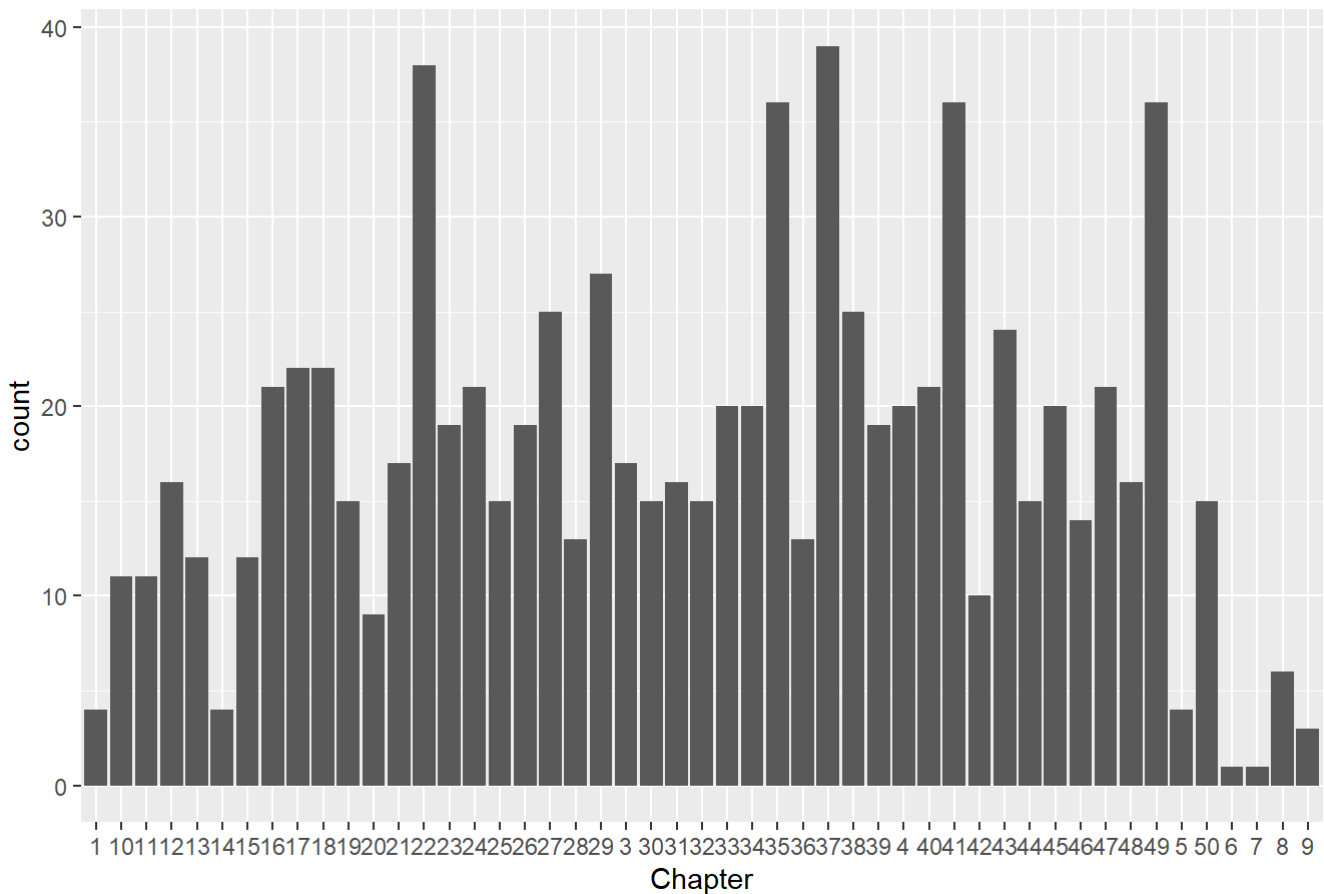
Elinor and Edward

```
ee <- tnum.query("@ref:EE", max = 1000)
```

```
## Returned 1 thru 851 of 851 results
```

```
eedf <- tnum.objectsToDf(ee)
ggplot()+geom_bar(mapping = aes(getchapter(ee)), stat = "count")+
  labs(x="Chapter", title = "Frequency of Elinor and Edward")
```

Frequency of Elinor and Edward



From this bar plot, we can see that the frequency of Elinor and Edward is fairly uniformly distributed in all the chapter, because this couple are in love with each other from the beginning to the end.

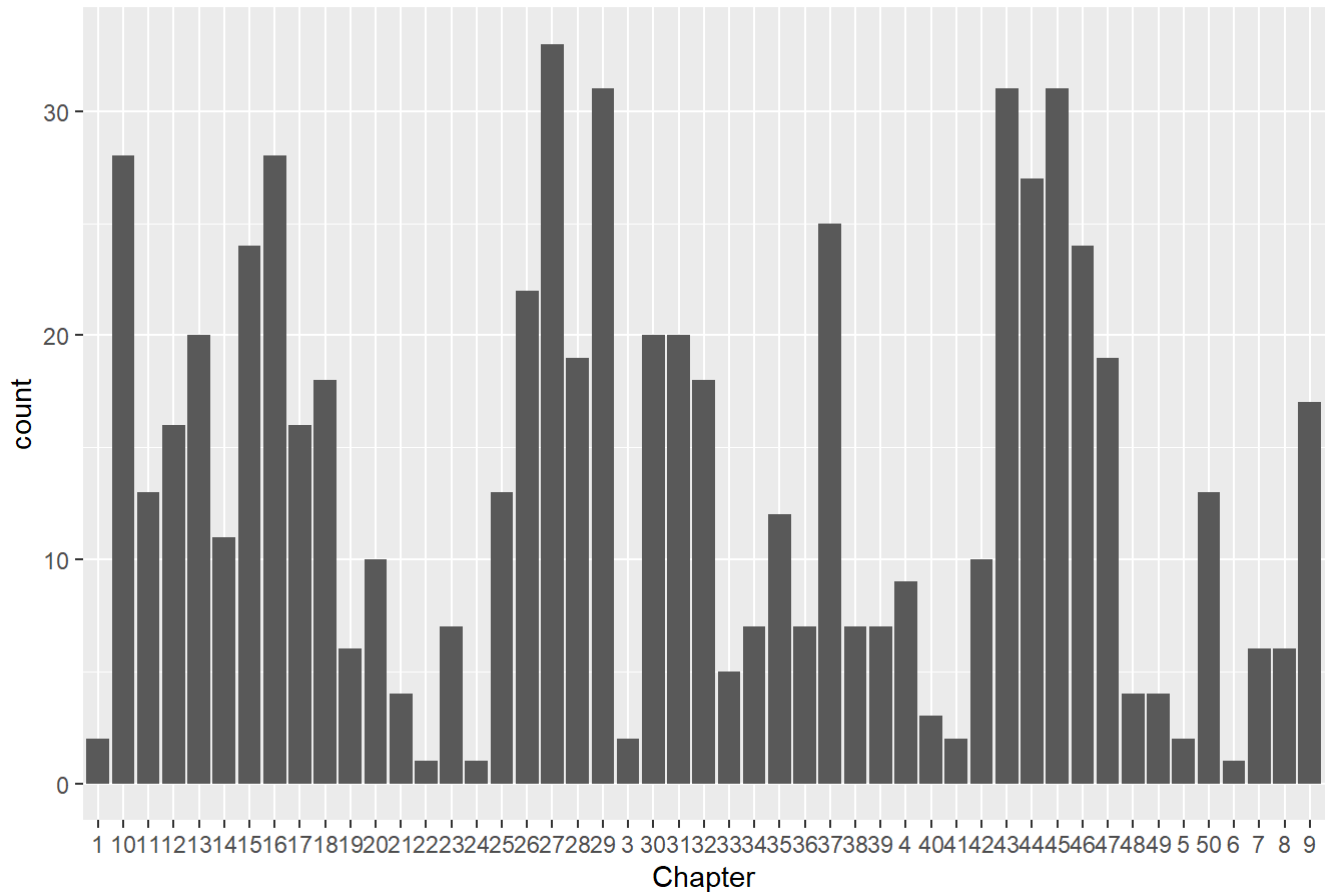
Marianne and Willoughby

```
mw <- tnum.query("@ref:MW", max = 1000)
```

```
## Returned 1 thru 662 of 662 results
```

```
mwdf <- tnum.objectsToDf(mw)
ggplot()+geom_bar(mapping = aes(getchapter(mw)), stat = "count")+
  labs(x="Chapter", title = "Frequency of 'Marianne' and 'Willoughby'")
```

Frequency of 'Marianne' and 'Willoughby'



From this plot, we can see that the appearance of both Marianne and Brandon is mainly in former 30 chapters in which they are in love. Although Willoughby married with others since chapter 30, he and Marianne still have an ups and downs of emotional relationship which matches the ups and downs in the latter part of the bar plot.

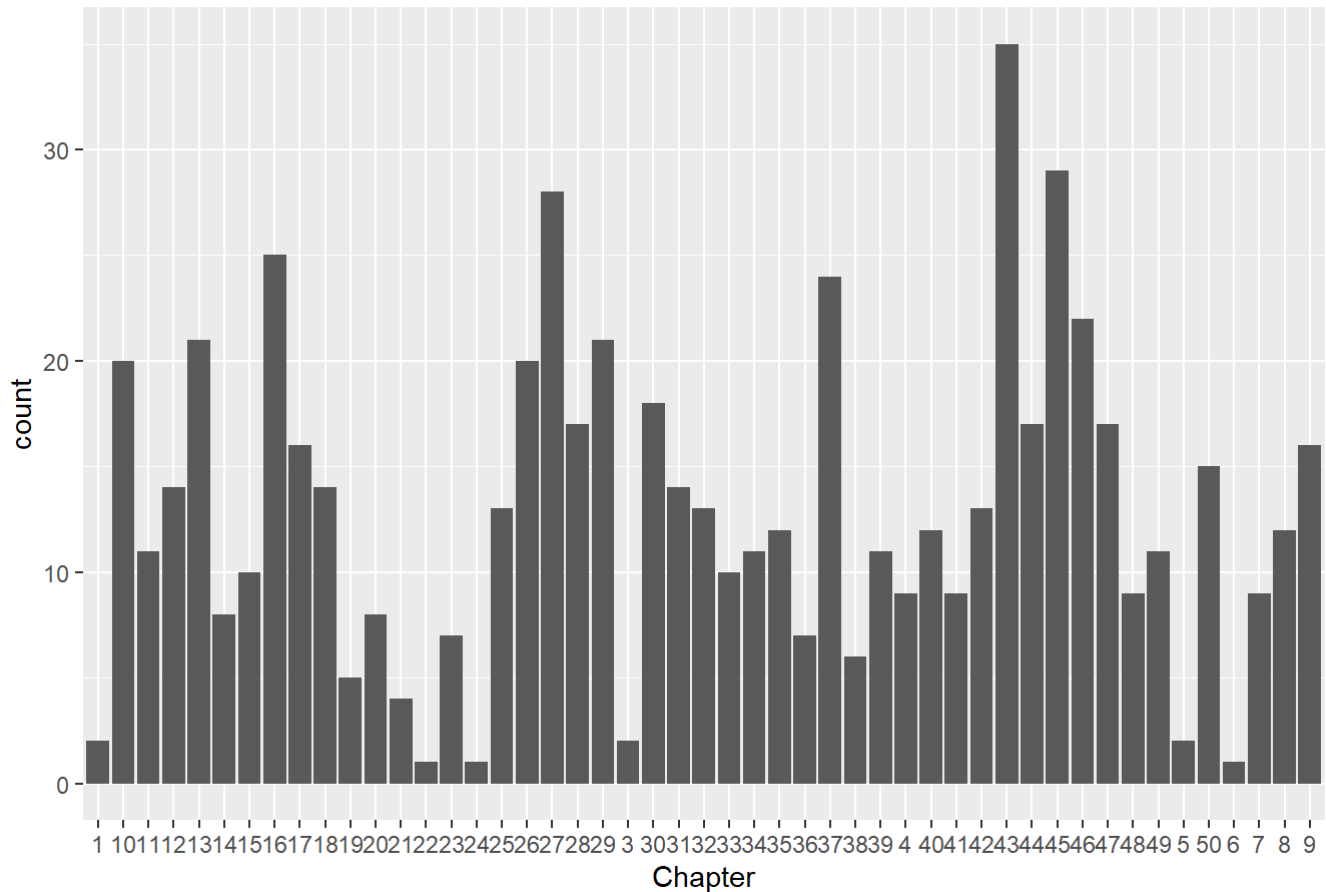
Marriane and Brandon

```
mb <- tnum.query("@ref:MB", max = 1500)
```

```
## Returned 1 thru 632 of 632 results
```

```
mbdf <- tnum.objectsToDf(mb)
ggplot()+geom_bar(mapping = aes(getchapter(mb)), stat = "count")+
  labs(x="Chapter", title = "Frequency of 'Marianne' and 'Brandon'")
```

Frequency of 'Marianne' and 'Brandon'



The appearance of Marianne and Brandon is concentrated on the later half of the book, especially after chapter 30. In the former chapters, Marianne is in love with Willoughby and Brandon is attracted by Elinor. But later, they begin to learn about each other and get married.

Couples with marriage

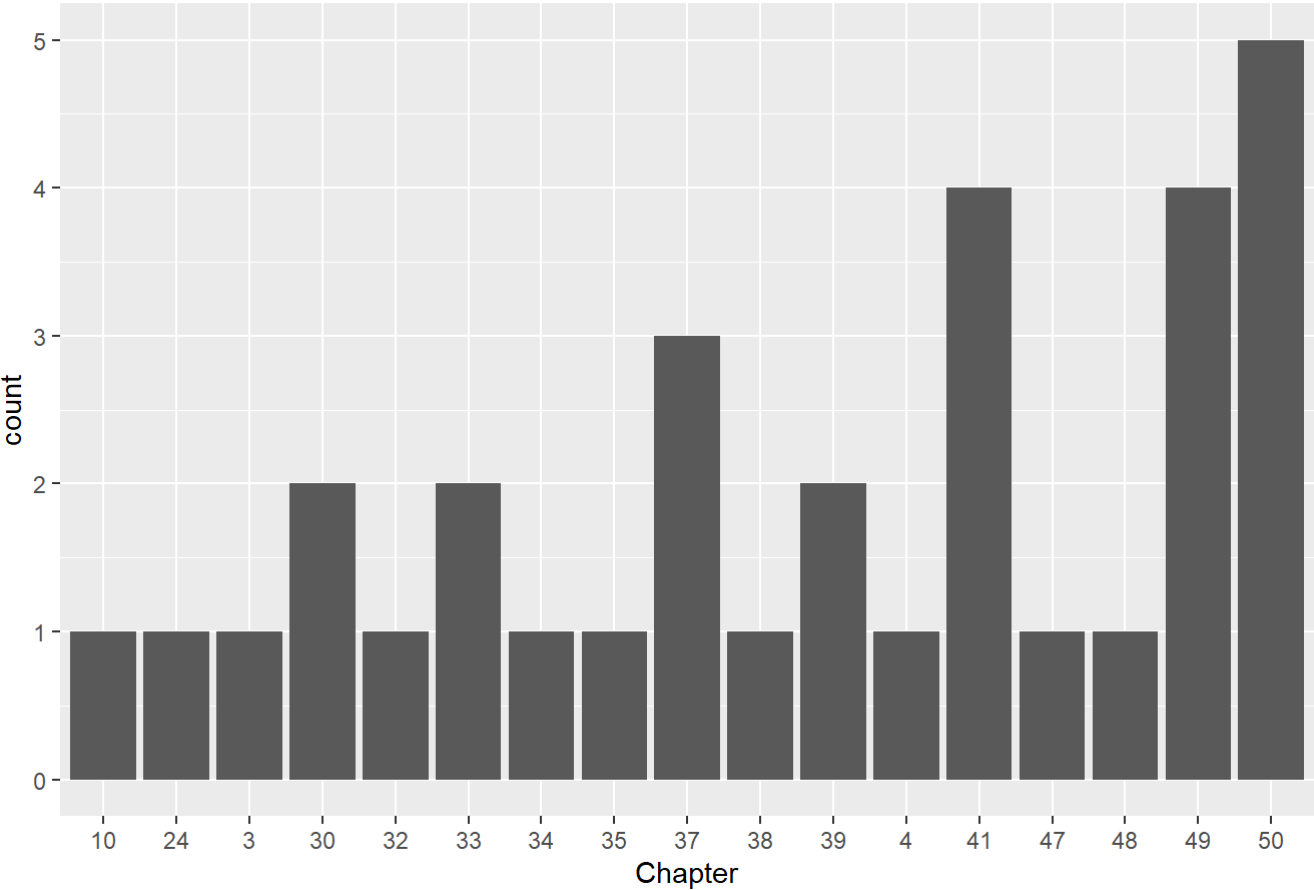
Now I want to view where both names and marriage appear so that we can learn about the relationship progress of these couples.

```
marryee <- tnum.query("@[ref:marriage,ref:EE]", max = 100)
```

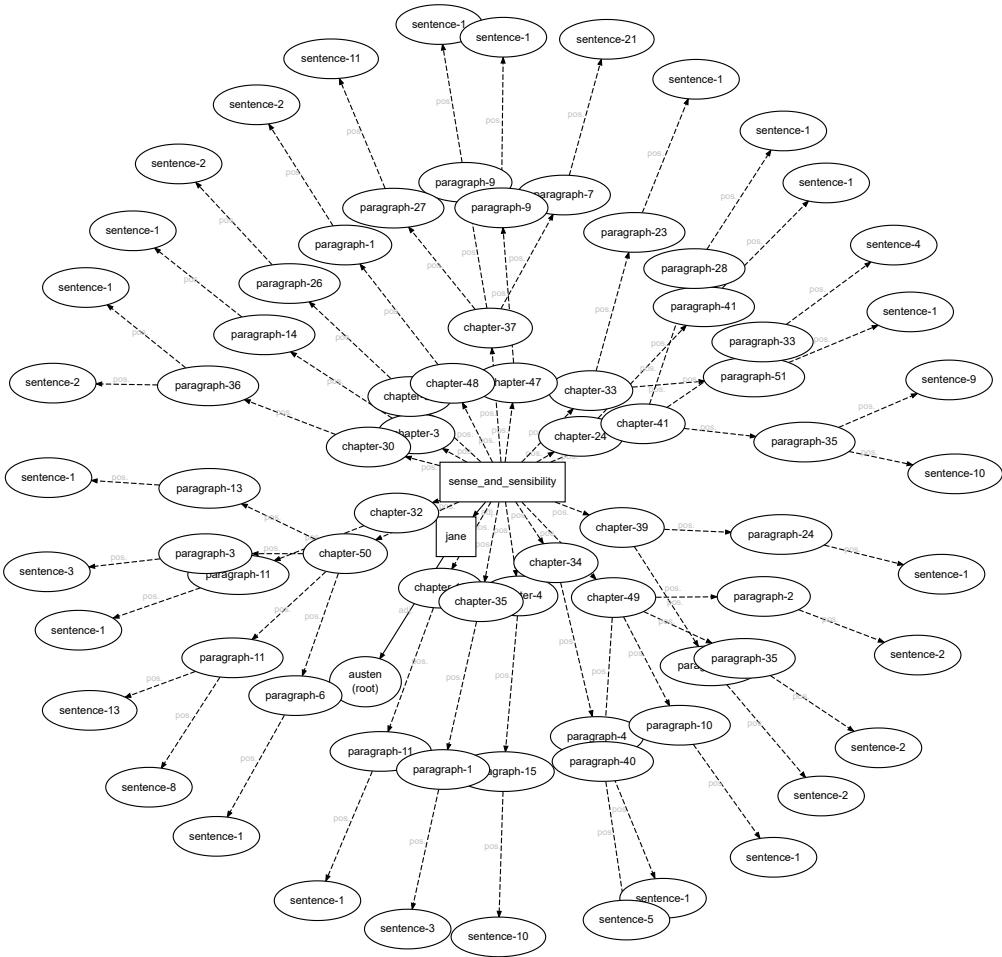
```
## Returned 1 thru 32 of 32 results
```

```
marryeef <- tnum.objectsToDf(marryee)
ggplot()+geom_bar(mapping = aes(getchapter(marryee)), stat = "count")+
  labs(x="Chapter", title = "Frequency of Elinor, Edward and marry")
```

Frequency of Elinor, Edward and marry



```
meegraph <- tnum.makePhraseGraphFromPathList(marryedf$subject)
tnum.plotGraph(meegraph)
```

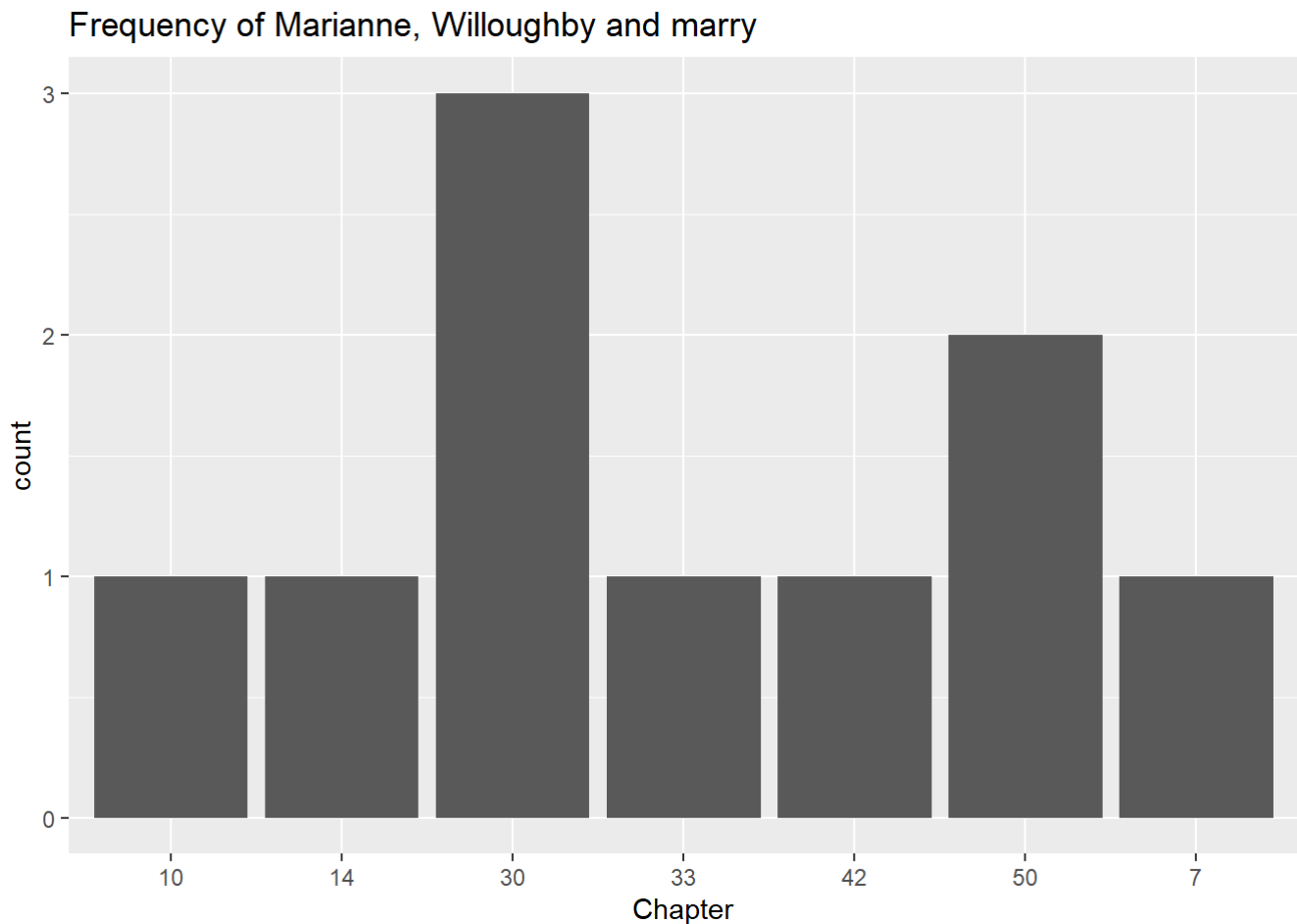


Elinor expresses her understanding of Edward's decision to marry with Lucy in chapter 37. Chapter 41 is a turning point of the relationship progress between Elinor and Edward because Lucy abandons the marriage appointment with Edward in this chapter. And in chapter 49-50, the couple is reconciled and get married.

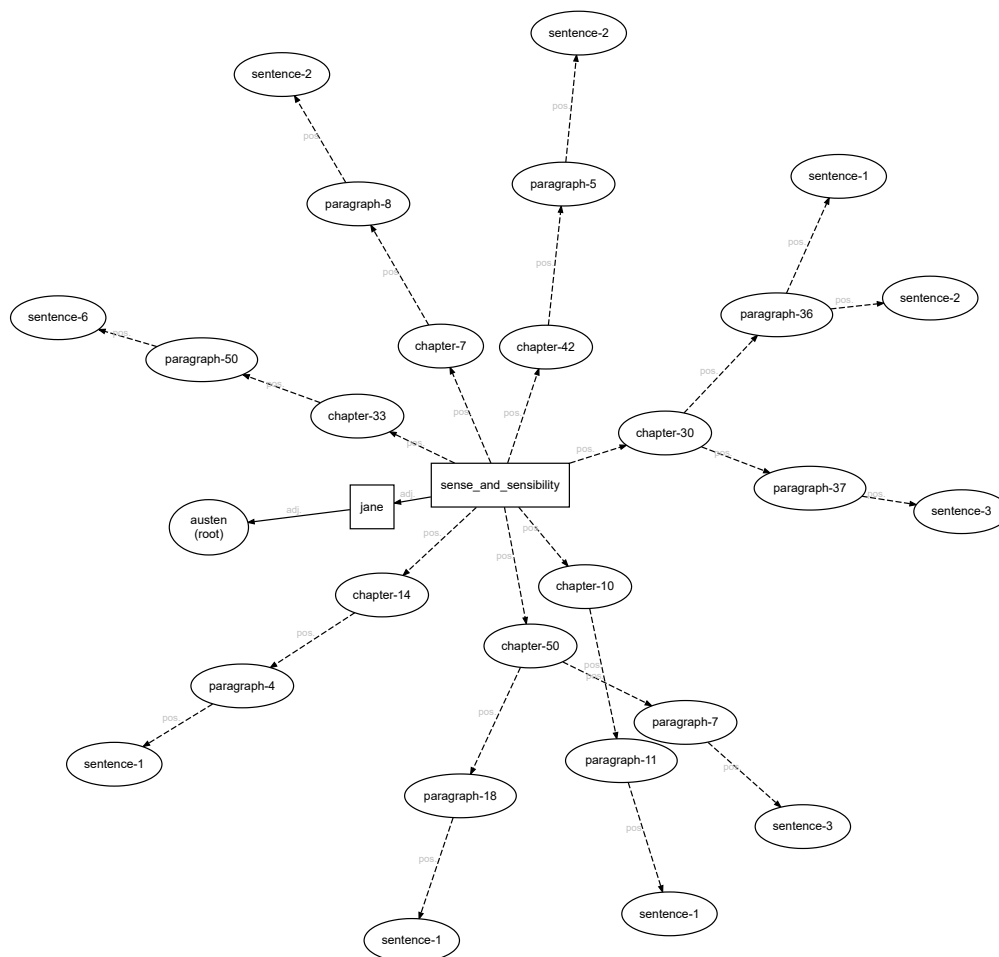
```
marrymw <- tnum.query("@[ref:marriage,ref:MW]", max = 100)
```

```
## Returned 1 thru 10 of 10 results
```

```
marrymwdf <- tnum.objectsToDf(marrymw)
ggplot()+geom_bar(mapping = aes(getchapter(marrymw)), stat = "count")+
  labs(x="Chapter", title = "Frequency of Marianne, Willoughby and marry")
```



```
mmwgraph <- tnum.makePhraseGraphFromPathList(marrymwdf$subject)
tnum.plotGraph(mmwgraph)
```



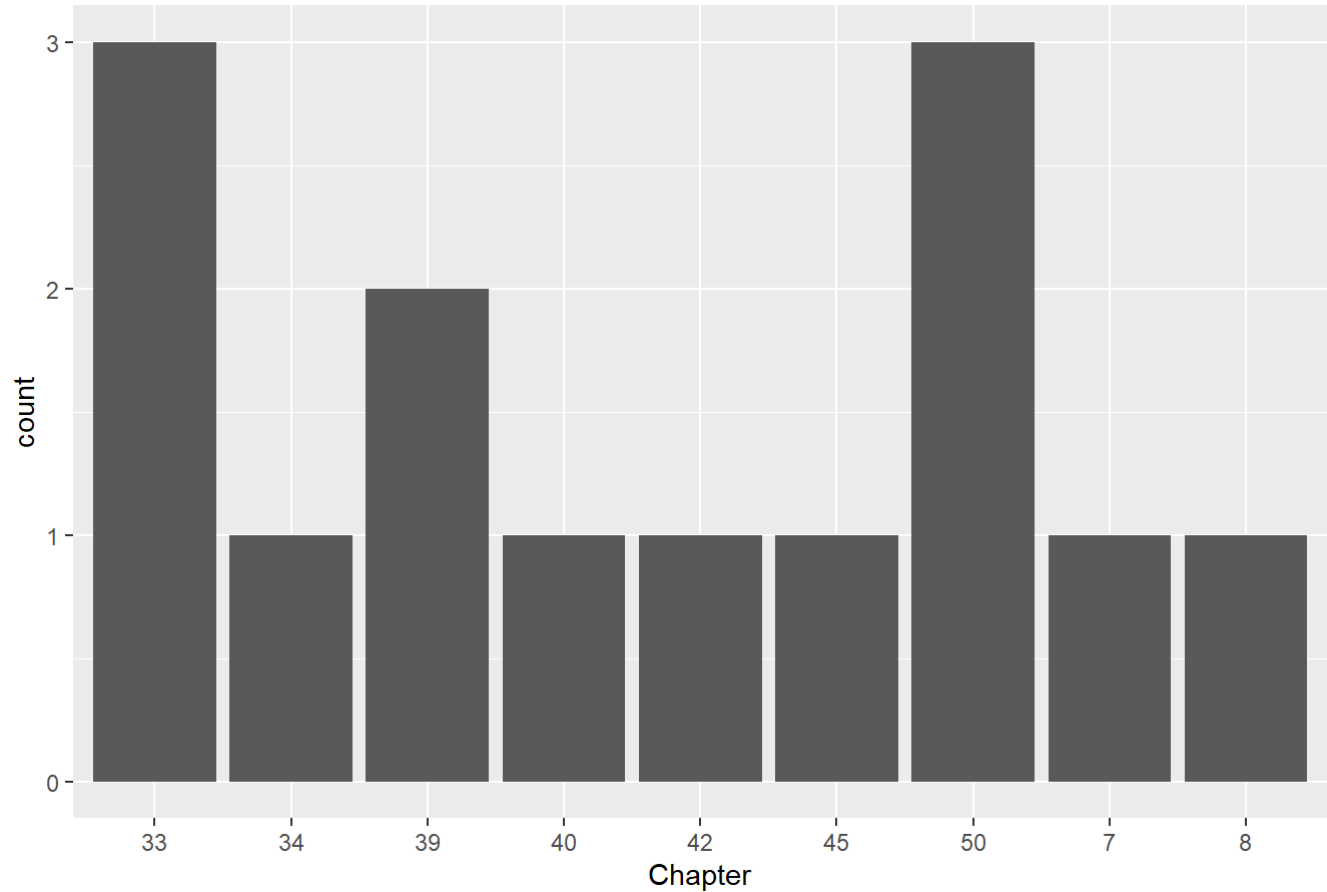
The frequency of Marianne, Willoughby and marriage are the largest in chapter 30. It's reasonable because Marianne hear of the marriage of Willoughby and another girl in this chapter which is a turning point of Marianne's emotional life.

```
marrymb <- tnum.query("@[ref:marriage,ref:MB]", max = 100)
```

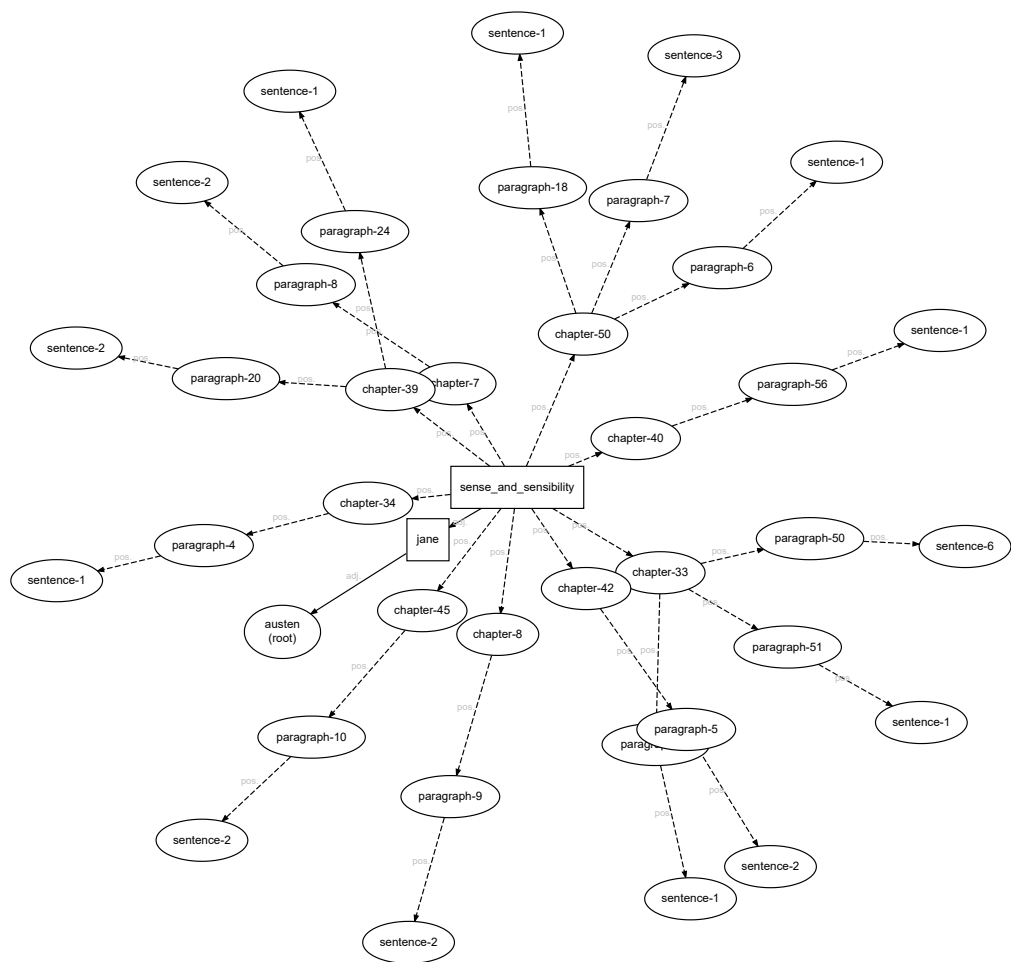
```
## Returned 1 thru 14 of 14 results
```

```
marrymbdf <- tnum.objectsToDf(marrymb)
ggplot()+geom_bar(mapping = aes(getchapter(marrymb)), stat = "count")+
  labs(x="Chapter", title = "Frequency of Marianne, Brandon and marry")
```


Frequency of Marianne, Brandon and marry



```
mmbgraph <- tnum.makePhraseGraphFromPathList(marrymbdf$subject)
tnum.plotGraph(mmbgraph)
```



Brandon steps into the emotional life of Marianne since chapter 33, which is just after Marianne was crossed in love in chapter 30. In chapter 50, Marianne accept the love of Brandon and get married with him.

Pride and Prejudice

Create tags

For *Pride and Prejudice* talks about different attitudes of girls towards marriage and love, there are marriages running through the whole book: Collins and Charlotte, Lydia and Wickham, Darcy and Elizabeth. So I create tags with marriage and marry, and the names of three couples separately.

```
tnum.tagByQuery("*pride* has*= REGEXP(\"marriage|marry\")", adds = ("ref:marry"))
```

```
## list(modifiedCount = 120, tagged = 120, removed = 0)
```

```
tnum.tagByQuery("*pride* has*= REGEXP(\"Collins|Charlotte\")", adds = ("ref:CC"))
```

```
## list(modifiedCount = 236, tagged = 236, removed = 0)
```

```
tnum.tagByQuery("*pride* has*= REGEXP(\"Lydia|Wickham\")", adds = ("ref:LW"))
```

```
## list(modifiedCount = 319, tagged = 319, removed = 0)
```

```
tnum.tagByQuery("*pride* has*= REGEXP(\"Elizabeth|Darcy\")", adds = ("ref:ED"))
```

```
## list(modifiedCount = 930, tagged = 930, removed = 0)
```

Plots

I want to know where those roles appear frequently, and I can match with plots to see the reason why their names in these chapters were always mentioned.

Marry and marriage

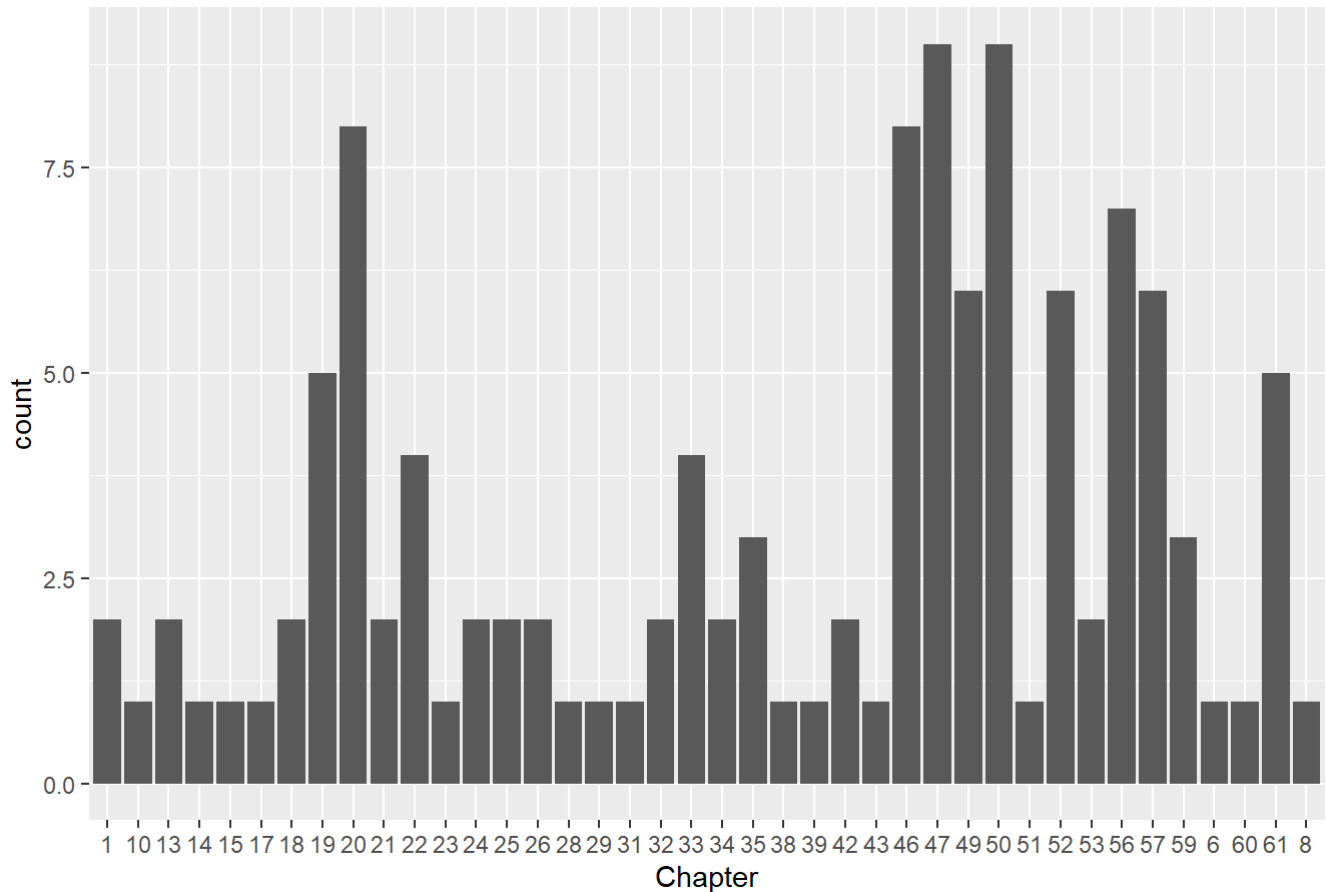
At first, to see in which chapters 'marriage' and 'marry' are mentioned frequently.

```
marry <- tnum.query("@ref:marry", max = 120)
```

```
## Returned 1 thru 120 of 239 results
```

```
marrydf <- tnum.objectsToDf(marry)
ggplot()+geom_bar(mapping = aes(getchapter(marry)), stat = "count")+
  labs(x="Chapter", title = "Frequency of 'marry' and 'marriage'")
```

Frequency of 'marry' and 'marriage'



It is clear that in chapter 19-22 and chapter 46-59, the frequency of these two words are outstanding. And if we trace back to the contents of *Pride and Prejudice*, in chapter 19-22, Collins and Charlotte got married; in chapter 46-49, Lydia and Wickham got married; in chapter 56-59, Darcy and Elizabeth got married.

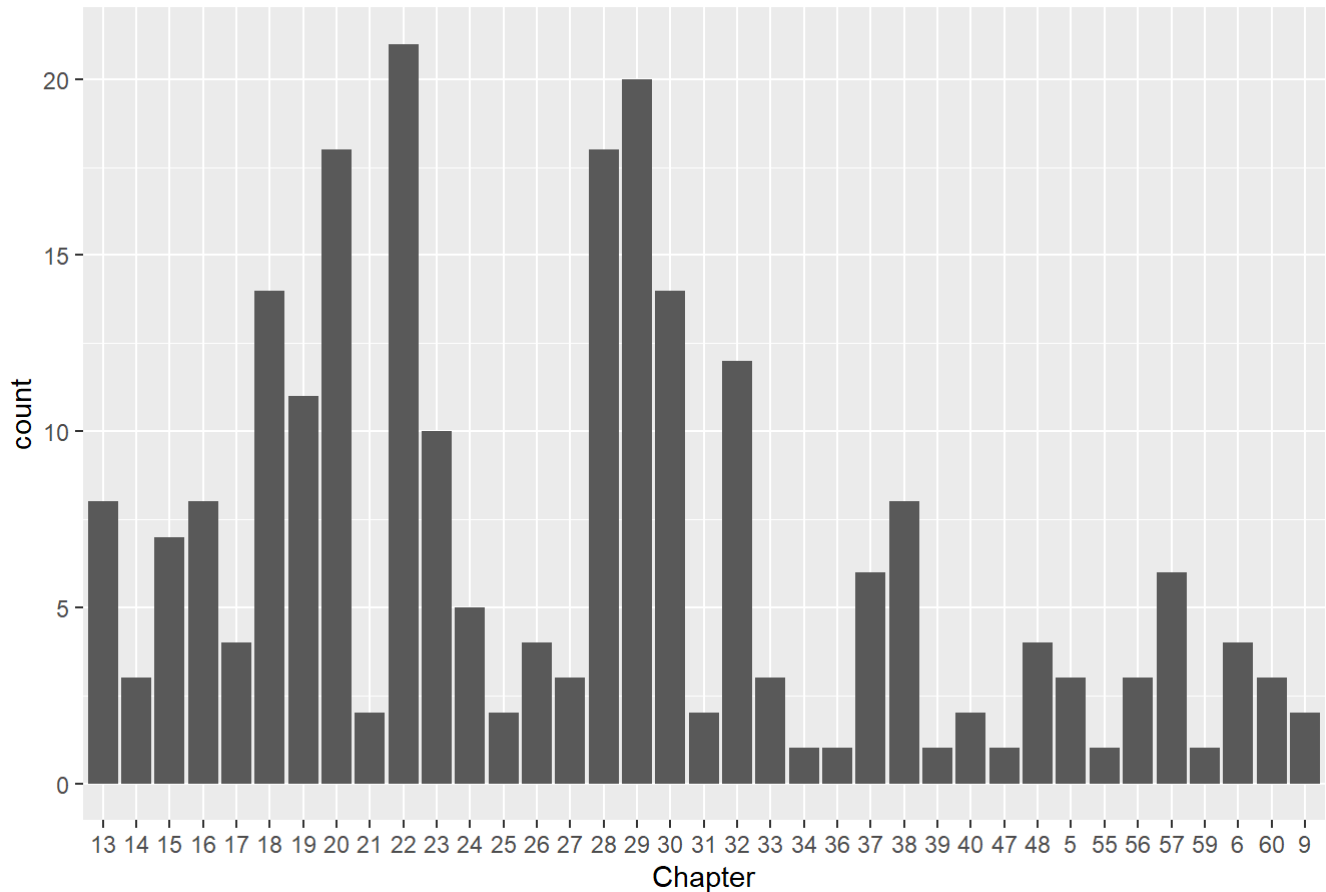
Collins and Charlotte

```
cc <- tnum.query("@ref:CC", max = 1000)
```

```
## Returned 1 thru 236 of 236 results
```

```
ccdf <- tnum.objectsToDf(cc)
ggplot()+geom_bar(mapping = aes(getchapter(cc)), stat = "count")+
  labs(x="Chapter", title = "Frequency of Collins and Charlotte")
```

Frequency of Collins and Charlotte



It is clear that Collins and Charlotte appear more frequently in the first half chapters. And in chapter 18-23 is the main stage of Collins and Charlotte, in chapter 28-30 they become frequently for Elizabeth came to visit them.

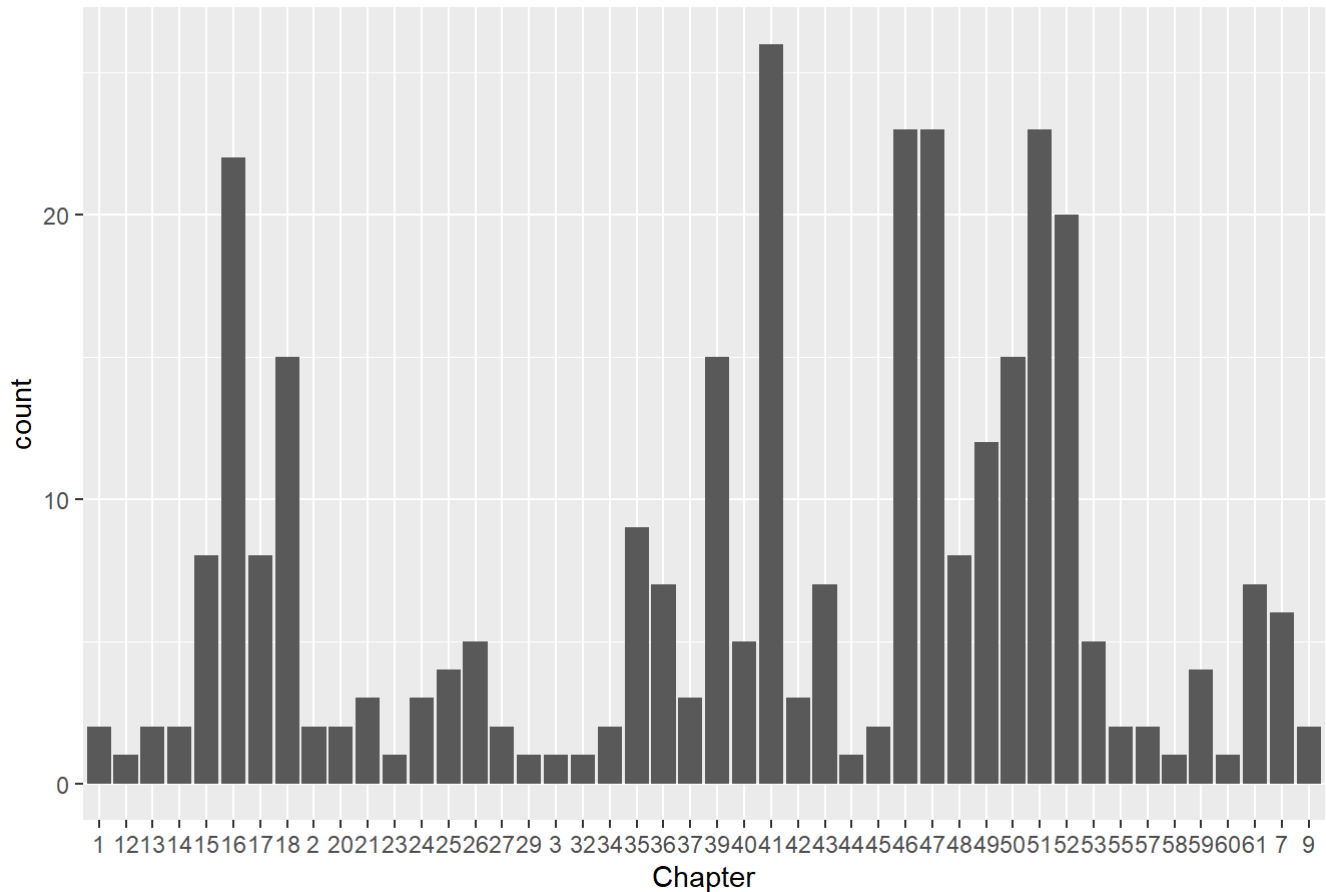
Lydia and Wickham

```
lw <- tnum.query("@ref:LW", max = 1000)
```

```
## Returned 1 thru 319 of 319 results
```

```
lwdf <- tnum.objectsToDf(lw)
ggplot()+geom_bar(mapping = aes(getchapter(lw)), stat = "count")+
  labs(x="Chapter", title = "Frequency of Lydia and Wickham")
```

Frequency of Lydia and Wickham



For Lydia and Wickham, the number of their names seems quite larger than Collins and Charlotte so they are more important roles to push the plots. The first meet of Elizabeth with Wickham is in chapter 15-18, and in chapter 35-41, Elizabeth knew the true side of Wickham, the marriage of Lydia and Wickham is in chapter 46-49.

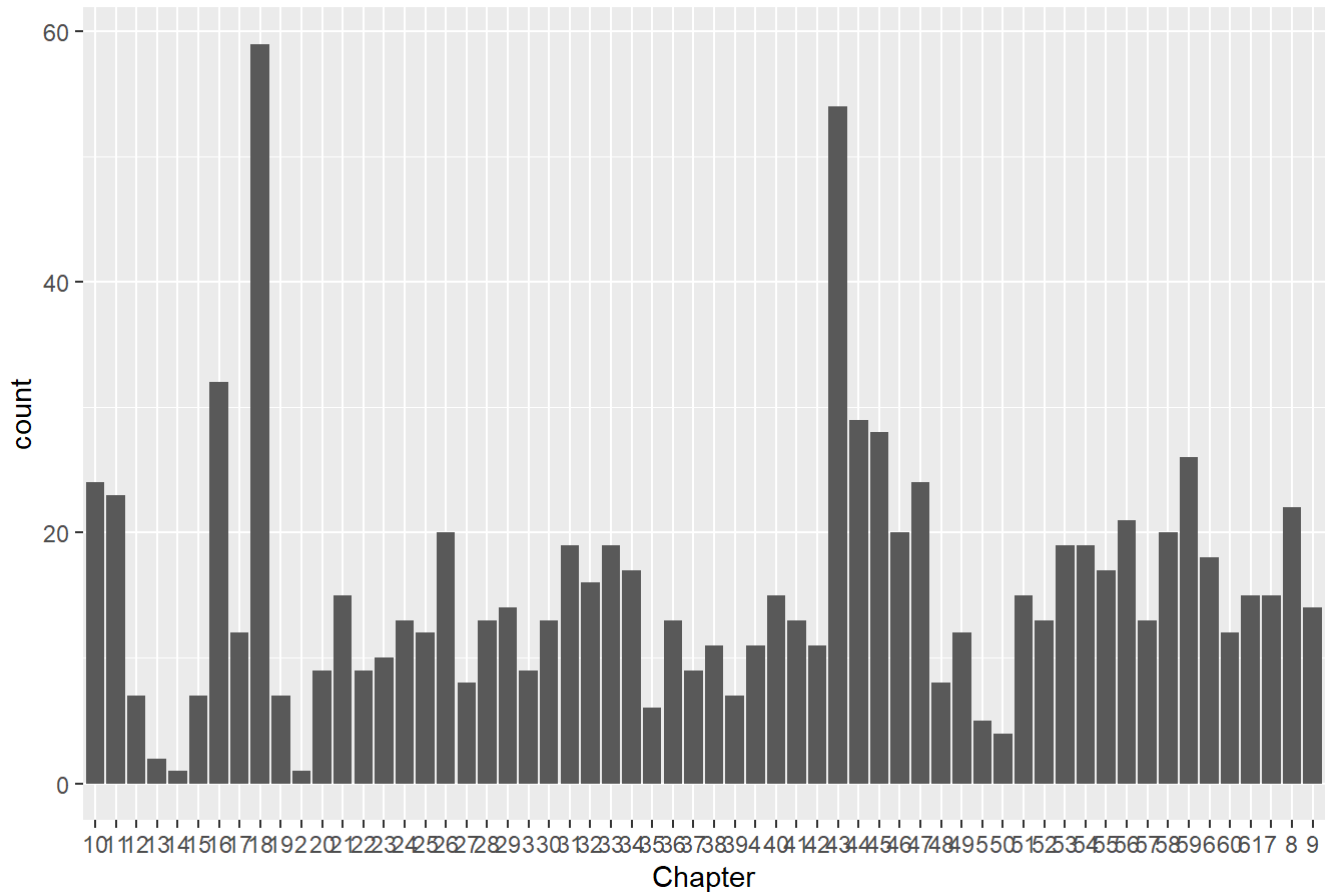
Darcy and Elizabeth

```
ed <- tnum.query("@ref:ED", max = 1000)
```

```
## Returned 1 thru 930 of 930 results
```

```
eddf <- tnum.objectsToDf(ed)
ggplot()+geom_bar(mapping = aes(getchapter(ed)), stat = "count")+
  labs(x="Chapter", title = "Frequency of Darcy and Elizabeth")
```

Frequency of Darcy and Elizabeth



Darcy and Elizabeth are the main characters in *Pride and Prejudice*, so, of course, the times of their names are much bigger. But there are still some outstanding chapters for they are not always in the same place. They first met in chapter 16-19, and they re-meet again in chapter 43-47.

Names of couples with marriage

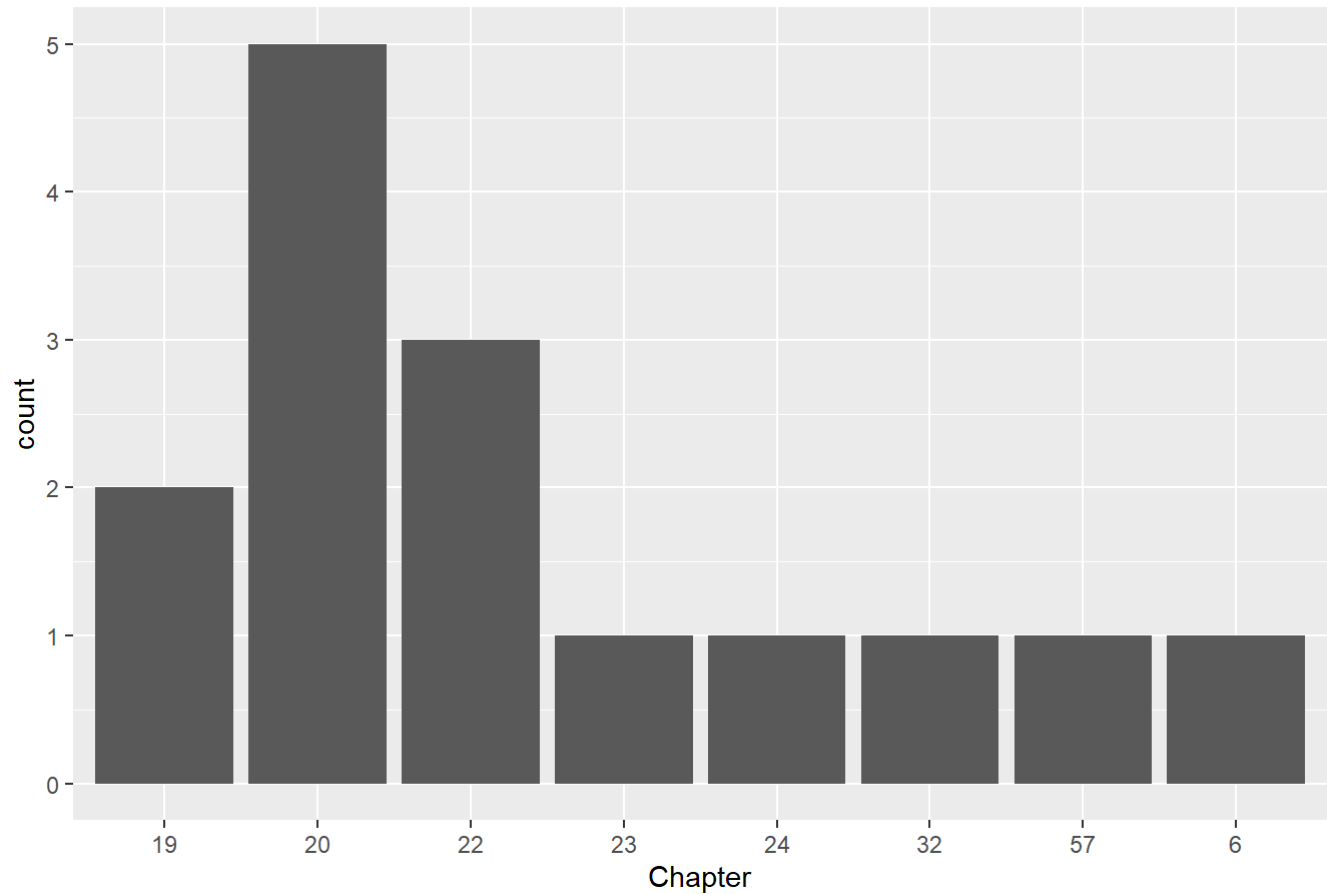
Secondly, I want to see where both names and marriage appear.

```
marrycc <- tnum.query("@[ref:marry,ref:CC]", max = 100)
```

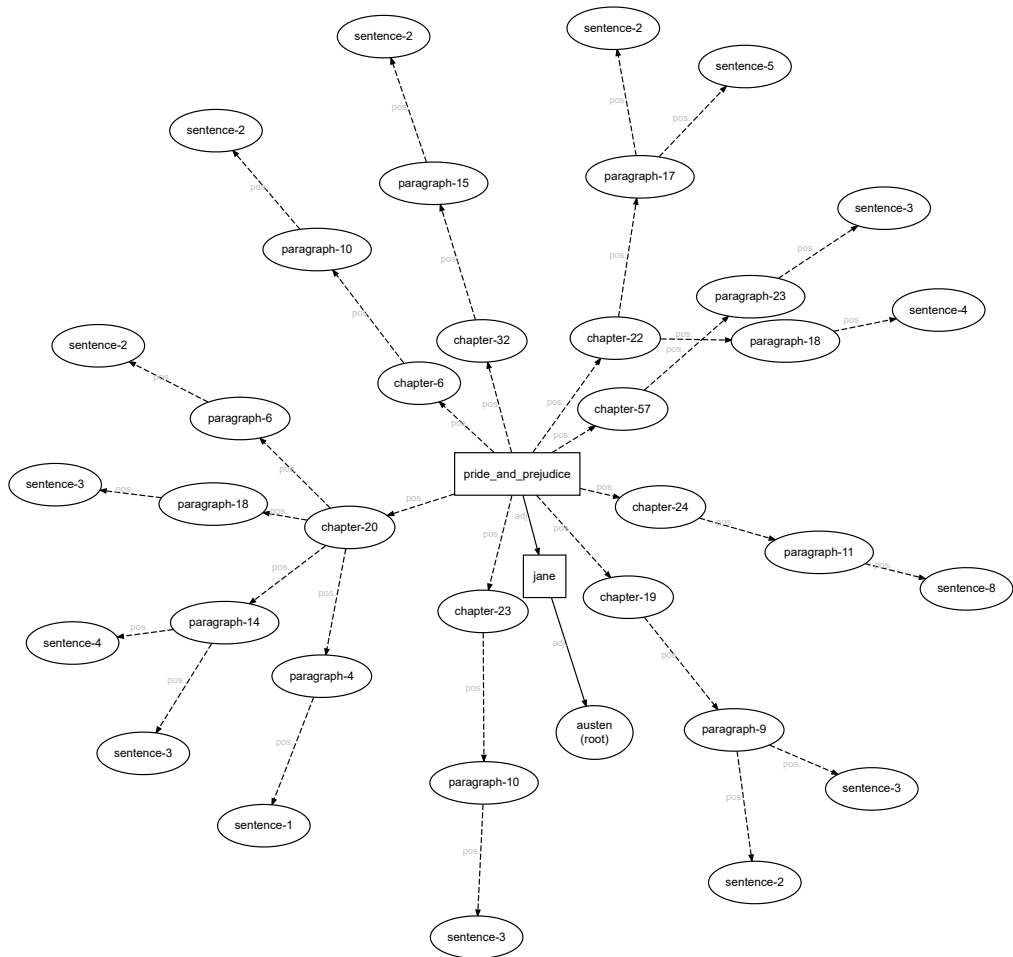
```
## Returned 1 thru 15 of 15 results
```

```
marryccdf <- tnum.objectsToDf(marrycc)
ggplot()+geom_bar(mapping = aes(getchapter(marrycc)), stat = "count")+
  labs(x="Chapter", title = "Frequency of Collins, Charlotte and marry")
```

Frequency of Collins, Charlotte and marry



```
mccgraph <- tnum.makePhraseGraphFromPathList(marryccdf$subject)
tnum.plotGraph(mccgraph)
```



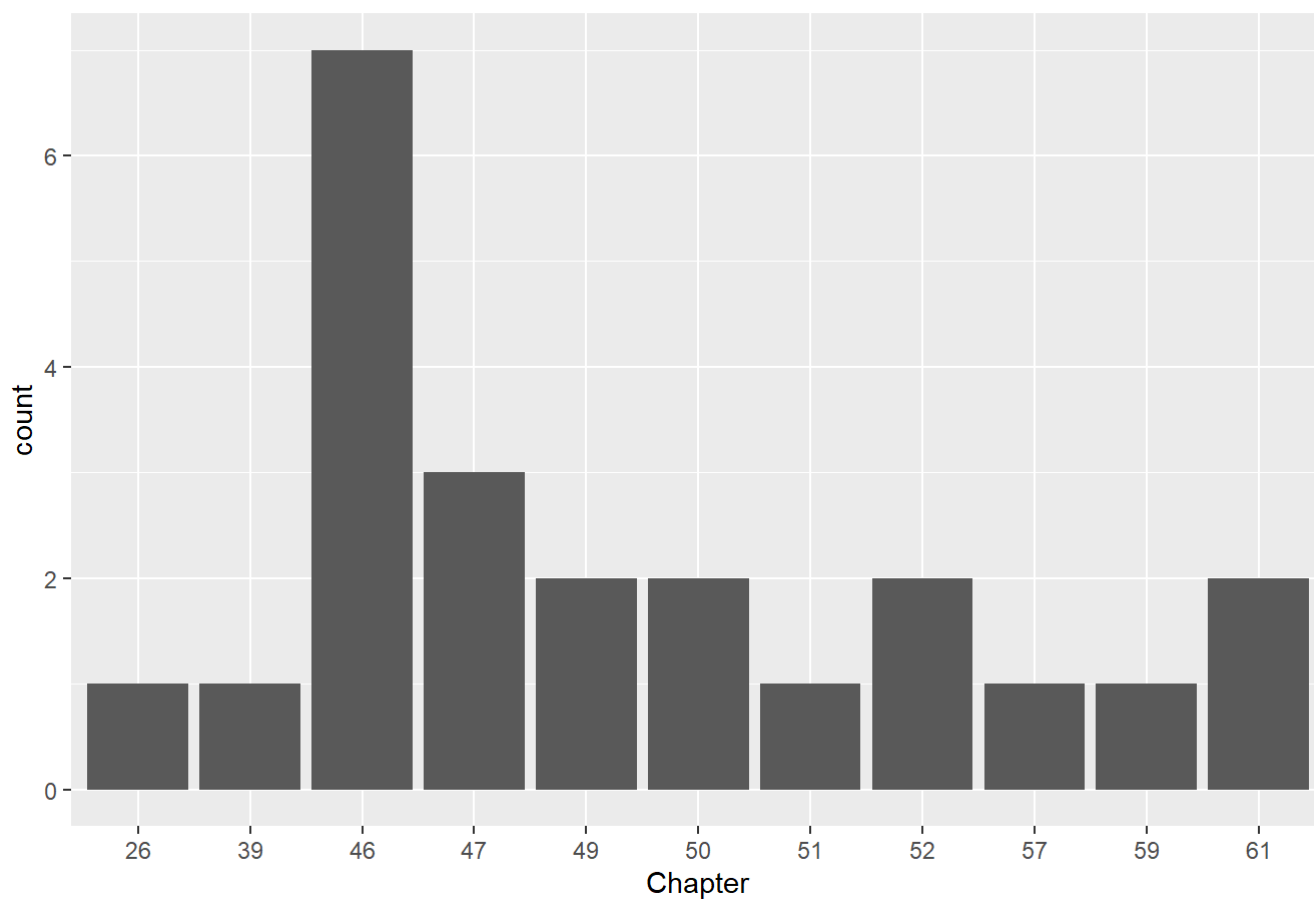
The barplot match the plot for Collins and Charlotte got married in chapter 19-22.

```
marrylw <- tnum.query("@[ref:marry,ref:LW]",max = 100)
```

```
## Returned 1 thru 23 of 23 results
```

```
marrylwdf <- tnum.objectsToDf(marrylw)
ggplot()+geom_bar(mapping = aes(getchapter(marrylw)),stat = "count")+
  labs(x="Chapter",title = "Frequency of Lydia, Wickham and marry")
```

Frequency of Lydia, Wickham and marry

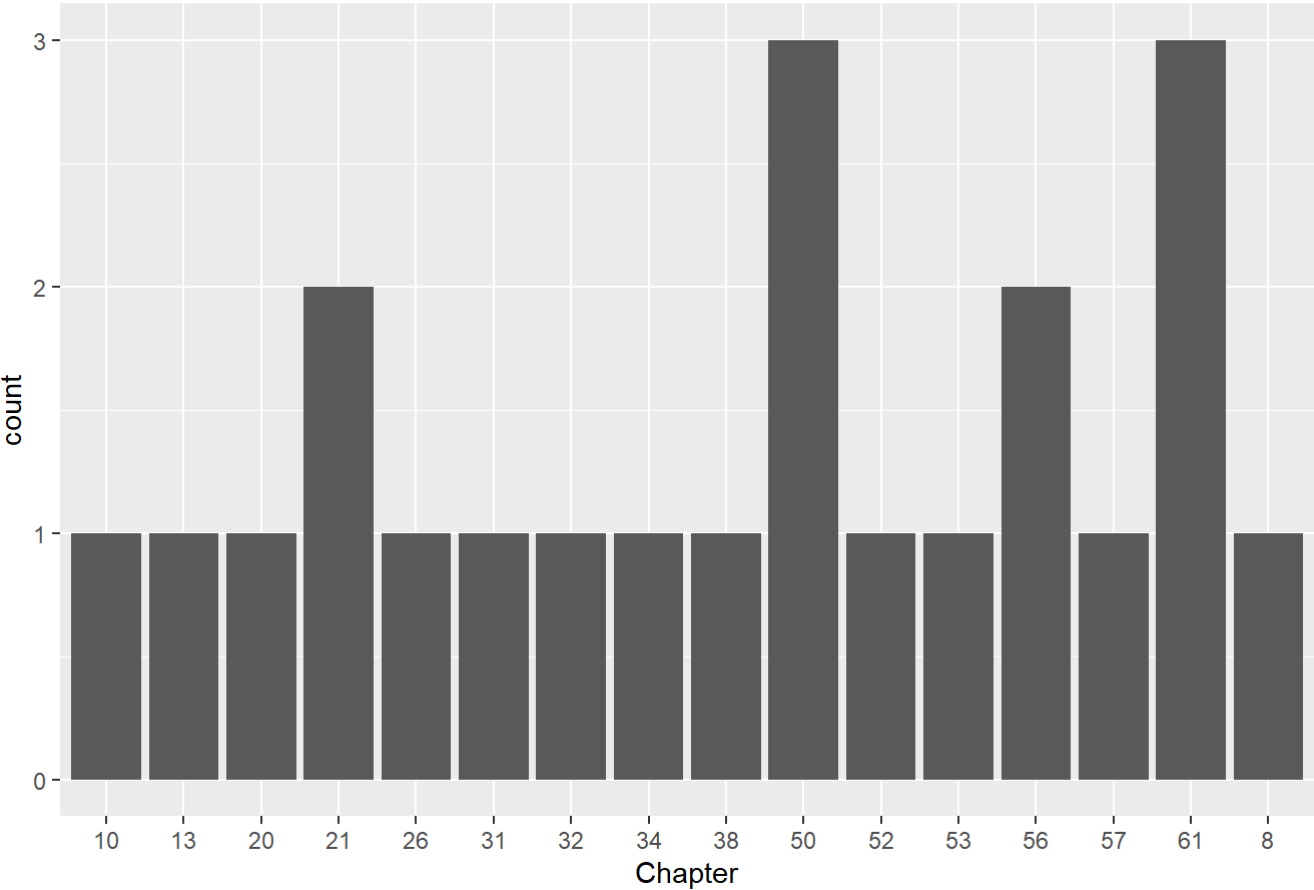


```
mlwgraph <- tnum.makePhraseGraphFromPathList(marrylwdf$subject)
tnum.plotGraph(mlwgraph)
```

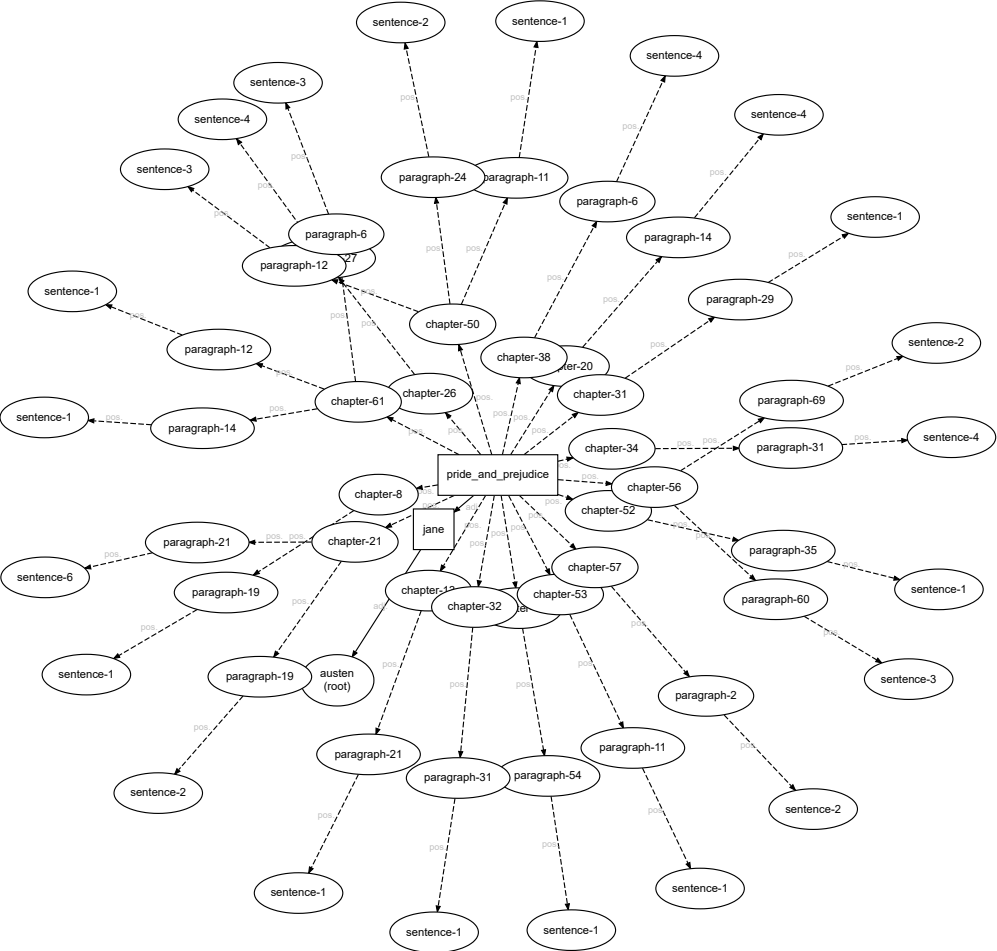



```
marrieddf <- tnum.objectsToDf(married)
ggplot()+geom_bar(mapping = aes(getchapter(married)), stat = "count")+
  labs(x="Chapter", title = "Frequency of Elizabeth, Darcy and marry")
```

Frequency of Elizabeth, Darcy and marry



```
medgraph <- tnum.makePhraseGraphFromPathList(marrieddf$subject)
tnum.plotGraph(medgraph)
```



For Darcy and Elizabeth the barplot does not give a clear point. But we can say that they got married late in around chapter 50-61.