Shakespeare Play Analysis

Import and Tidy Data

```
# Import Six Shakespeare Acts' Data
shakespeare = read_csv("shakespeare.csv")

# Pipe the shakespeare data frame to the next line
shakespeare %>%
  # Use count to find out how many titles/types there are
  count(title,type)
```

title <chr>	type <chr>	n <int>
A Midsummer Night's Dream	Comedy	3459
Hamlet, Prince of Denmark	Tragedy	6776
Much Ado about Nothing	Comedy	3799
The Merchant of Venice	Comedy	4225
The Tragedy of Macbeth	Tragedy	3188
The Tragedy of Romeo and Juliet	Tragedy	4441

6 rows

```
tidy_shakespeare <- shakespeare %>%
  # Group by the titles of the plays
  group_by(title) %>%
  # Define a new column linenumber
  mutate(linenumber = row_number()) %>%
  # Transform the non-tidy text data to tidy text data
  unnest_tokens(word, text) %>%
  ungroup()
```

Plot the Cleaned Shakespeare Data

```
# Plot Sentiment Changes with Respect to PLays
tidy_shakespeare %>%
  inner_join(get_sentiments("bing")) %>%
  count(title, type, index = linenumber %/% 70, sentiment) %>%
  spread(sentiment, n, fill = 0) %>%
  mutate(sentiment = positive - negative,
         fixed_title = str_wrap(title, width = 20)) %>%
  # Put index on x-axis, sentiment on y-axis, and map comedy/tragedy to fill
  ggplot(aes(x = index, y = sentiment, fill = type)) +
  # Make a bar chart with geom_col()
  geom_col() +
  # Separate panels for each title with facet_wrap()
  facet_wrap(~ fixed_title, scales = "free_x") +
  labs(x = "Index of Shakespeare Plays by 70 Lines Each",
       y = "Sentiment of Each Plays") +
  scale_fill_discrete(name = "Type")
```