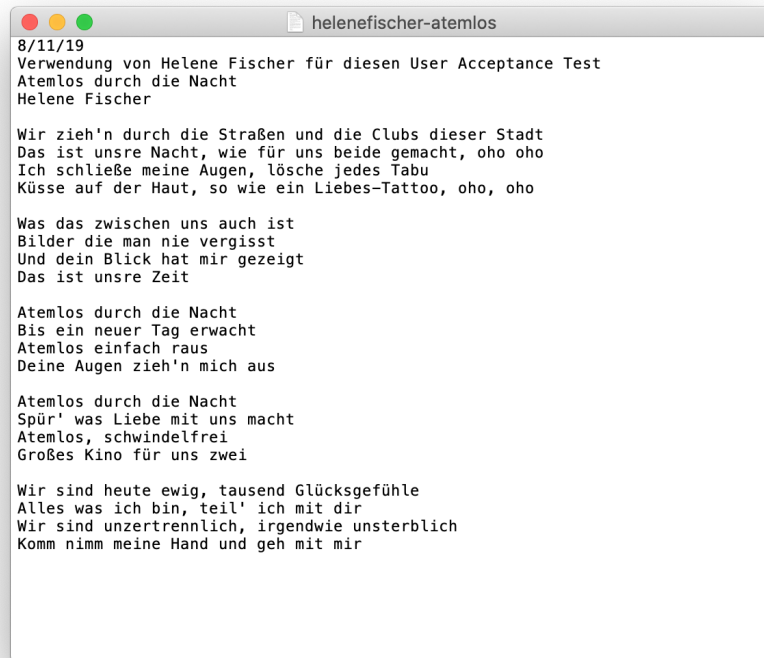# 1 User Acceptance Testing

8/11/19 - 10:50am

For this User Acceptance Test (UAT) to pass it will need to:

1. Extract body text from a .txt file and create a new .txt file with that text.

2. Extract title text from a .txt file and create a new .txt file with that text.

3. Extract title and subtitle text together from a .txt file and create a new .txt file with that text.

4. Analyse the body text for word frequency and create a .csv file that has that information.

5. Analyse the title text for word frequency and create a .csv file that has that information.

6. Analyse the title and subtitle text for word frequency and create a .csv file that has that information.

I have created a new .txt file for this UAT.

```
● ● ●                    📄 helenefischer-atemlos

8/11/19
Verwendung von Helene Fischer für diesen User Acceptance Test
Atemlos durch die Nacht
Helene Fischer

Wir zieh'n durch die Straßen und die Clubs dieser Stadt
Das ist unsre Nacht, wie für uns beide gemacht, oho oho
Ich schließe meine Augen, lösche jedes Tabu
Küsse auf der Haut, so wie ein Liebes-Tattoo, oho, oho

Was das zwischen uns auch ist
Bilder die man nie vergisst
Und dein Blick hat mir gezeigt
Das ist unsre Zeit

Atemlos durch die Nacht
Bis ein neuer Tag erwacht
Atemlos einfach raus
Deine Augen zieh'n mich aus

Atemlos durch die Nacht
Spür' was Liebe mit uns macht
Atemlos, schwindelfrei
Großes Kino für uns zwei

Wir sind heute ewig, tausend Glücksgefühle
Alles was ich bin, teil' ich mit dir
Wir sind unzertrennlich, irgendwie unsterblich
Komm nimm meine Hand und geh mit mir
```

I have also duplicated it 4 times so there are 5 copies of the same .txt file. This is to test if the Proof of Concept can accept multiple .txt files as a corpus.
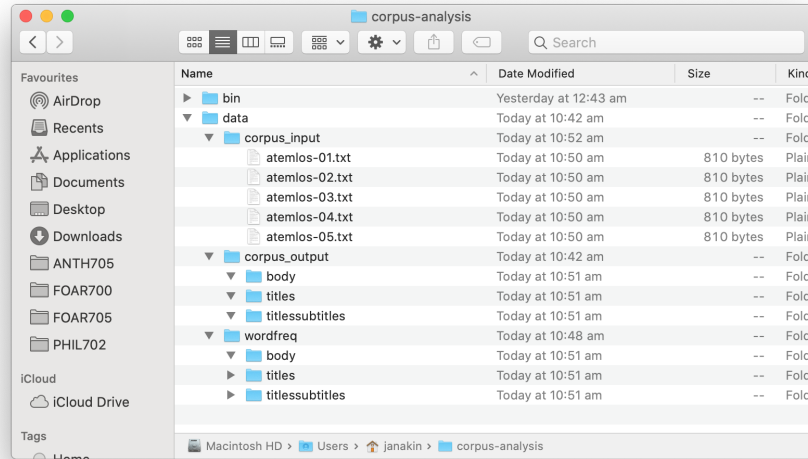
Having counted the frequency of words in the body text lyrics I will use, I expect that it would yield the same results multiplied by 5.

- atemlos - 20
- oho - 20
- nacht - 15
- ziehn - 10
- unsre - 10
- augen - 10

For the titles and titles and subtitles result I should expect no word over 5 as they only occur once.

Titles

- verwendung - 5
- helene - 5
- fischer - 5
- user - 5
- acceptance - 5

- test - 5

Titles and Subtitles

- verwendung - 5
- helene - 5
- fischer - 5
- user - 5
- acceptance - 5
- test - 5
- atemlos - 5
- nacht - 5

Now for the test.

Objective: Run UAT to see if Proof of Concept meets the Objective Criteria.

Action: Opened Terminal, navigated to the bin directory and called:

```
bash go.sh
```

Error: None as such, but some warnings.

Result: This is what Terminal returned

```
Janakins-MBP:~ janakin$ cd corpus-analysis/bin
Janakins-MBP:bin janakin$ ls
delete.sh go.sh
extractbody.sh wordfreq_body.R
extracttitles.sh wordfreq_titles.R
extracttitlessubtitles.sh wordfreq_titlessubtitles.R
Janakins-MBP:bin janakin$ bash go.sh
Loading required package: tm
Loading required package: NLP
Warning message:
In tm_map.SimpleCorpus(VCorpus, stripWhitespace) :
  transformation drops documents
Warning message:
In tm_map.SimpleCorpus(VCorpus_clean, content_transformer(tolower)) :
```

```
    transformation drops documents
Warning message:
In tm_map.SimpleCorpus(VCorpus_clean, removeWords, stopwords("german")) :
  transformation drops documents
Warning message:
In tm_map.SimpleCorpus(VCorpus_clean, removePunctuation) :
  transformation drops documents
Warning message:
In tm_map.SimpleCorpus(VCorpus_clean, removeNumbers) :
  transformation drops documents
Warning message:
In tm_map.SimpleCorpus(VCorpus_clean, removeWords, c("", "|")) :
  transformation drops documents
Loading required package: tm
Loading required package: NLP
Warning message:
In tm_map.SimpleCorpus(VCorpus, stripWhitespace) :
  transformation drops documents
Warning message:
In tm_map.SimpleCorpus(VCorpus_clean, content_transformer(tolower)) :
  transformation drops documents
Warning message:
In tm_map.SimpleCorpus(VCorpus_clean, removeWords, stopwords("german")) :
  transformation drops documents
Warning message:
In tm_map.SimpleCorpus(VCorpus_clean, removePunctuation) :
  transformation drops documents
Warning message:
In tm_map.SimpleCorpus(VCorpus_clean, removeNumbers) :
  transformation drops documents
Warning message:
In tm_map.SimpleCorpus(VCorpus_clean, removeWords, c("", "|")) :
  transformation drops documents
Loading required package: tm
Loading required package: NLP
Warning message:
In tm_map.SimpleCorpus(VCorpus, stripWhitespace) :
  transformation drops documents
Warning message:
In tm_map.SimpleCorpus(VCorpus_clean, content_transformer(tolower)) :
  transformation drops documents
Warning message:
In tm_map.SimpleCorpus(VCorpus_clean, removeWords, stopwords("german")) :
  transformation drops documents
Warning message:
In tm_map.SimpleCorpus(VCorpus_clean, removePunctuation) :
```
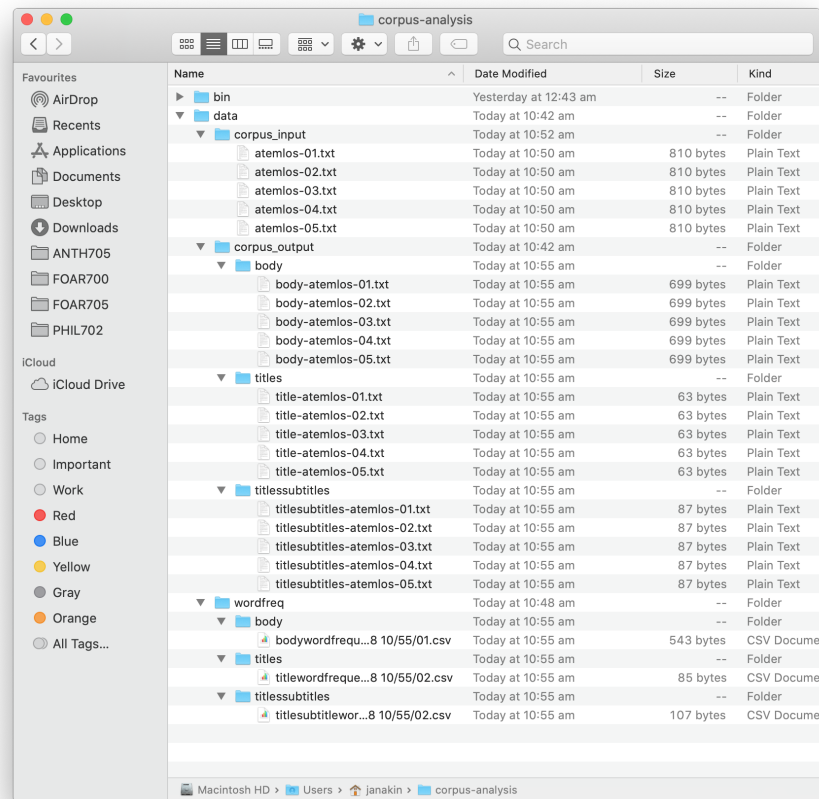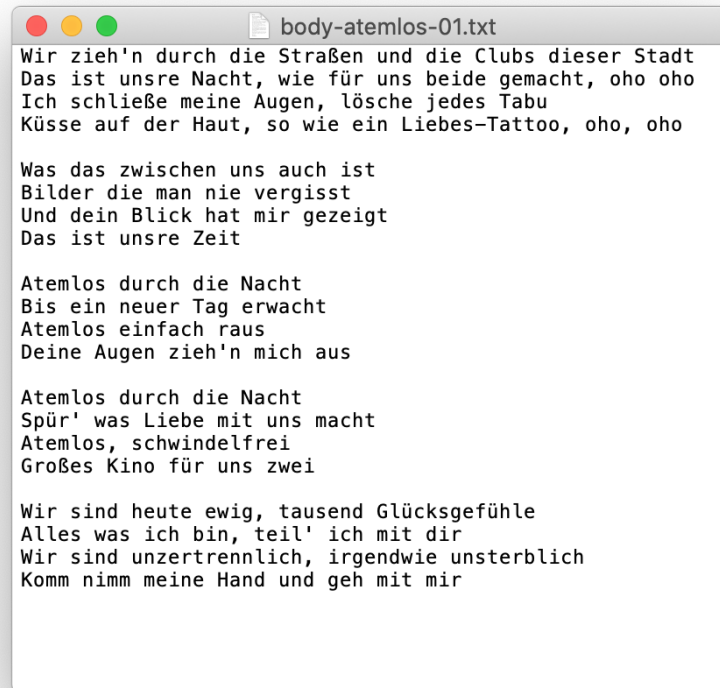
```
        transformation drops documents
Warning message:
In tm_map.SimpleCorpus(VCorpus_clean, removeNumbers) :
  transformation drops documents
Warning message:
In tm_map.SimpleCorpus(VCorpus_clean, removeWords, c("", "|")) :
  transformation drops documents
Janakins-MBP:bin janakin$
```
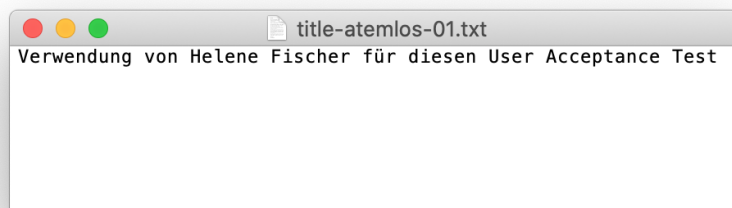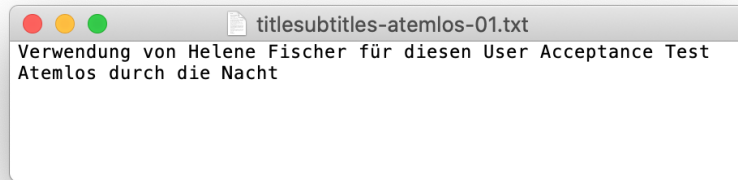
The files were created in the right areas.

The body text was extracted correctly.

```
● ● ●                    📄 body-atemlos-01.txt
Wir zieh'n durch die Straßen und die Clubs dieser Stadt
Das ist unsre Nacht, wie für uns beide gemacht, oho oho
Ich schließe meine Augen, lösche jedes Tabu
Küsse auf der Haut, so wie ein Liebes-Tattoo, oho, oho

Was das zwischen uns auch ist
Bilder die man nie vergisst
Und dein Blick hat mir gezeigt
Das ist unsre Zeit

Atemlos durch die Nacht
Bis ein neuer Tag erwacht
Atemlos einfach raus
Deine Augen zieh'n mich aus

Atemlos durch die Nacht
Spür' was Liebe mit uns macht
Atemlos, schwindelfrei
Großes Kino für uns zwei

Wir sind heute ewig, tausend Glücksgefühle
Alles was ich bin, teil' ich mit dir
Wir sind unzertrennlich, irgendwie unsterblich
Komm nimm meine Hand und geh mit mir
```

The title text was extracted correctly.

```
● ● ●                    📄 title-atemlos-01.txt
Verwendung von Helene Fischer für diesen User Acceptance Test
```

The title and subtitle text was extracted correctly.

```
  ● ● ●                     📄 titlesubtitles-atemlos-01.txt
Verwendung von Helene Fischer für diesen User Acceptance Test
Atemlos durch die Nacht
```

Results from the body text frequency analysis are below.

View	Zoom	Add Category	Insert	Table	Graph	Text	Format	Organise

+	Sheet 1

bodywordfrequenc
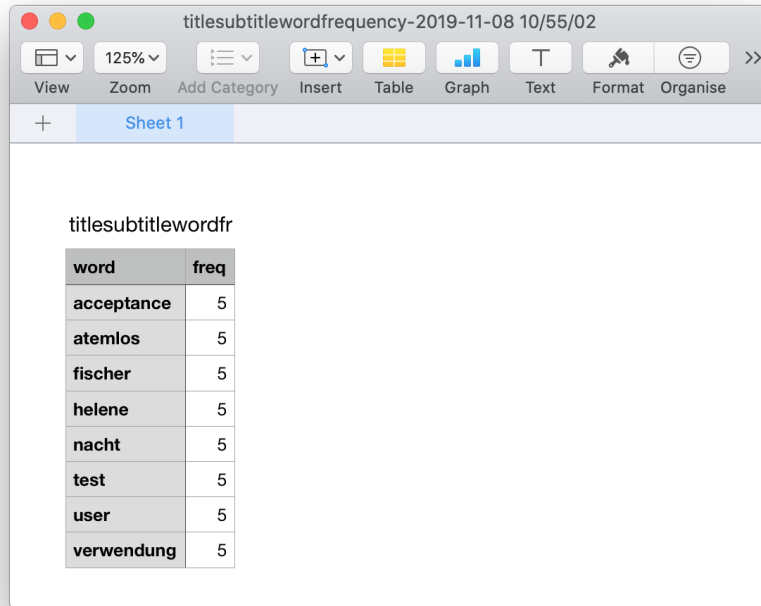
word	freq
atemlos	20
oho	20
nacht	15
augen	10
unsre	10
ziehn	10
beide	5
bilder	5
blick	5
clubs	5
einfach	5
erwacht	5
ewig	5
geh	5
gemacht	5
gezeigt	5
glücksgefühle	5
großes	5
hand	5
haut	5
heute	5
irgendwie	5
kino	5
komm	5
küsse	5
liebe	5
liebestattoo	5
lösche	5
macht	5

Results from the title text frequency analysis are below.

titlewordfrequency-2019-11-08 10/55/02

View	Zoom	Add Category
Insert	Table	Graph
Text	Format	Organise

Sheet 1

titlewordfrequenc

	word	freq
1	word	freq
2	acceptance	5
3	fischer	5
4	helene	5
5	test	5
6	user	5
7	verwendung	5

Results from the title and subtitle text frequency analysis are below.

The UAT Criteria have been met. The UAT has been passed.