Analysis of Vettius Valens

Definitions

```
library(ggplot2)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(cowplot)
```

```
##
## ************************************************************
```

```
## Note: As of version 1.0.0, cowplot does not change the
```

```
##   default ggplot2 theme anymore. To recover the previous
```

```
##   behavior, execute:
##   theme_set(theme_cowplot())
```

```
## ************************************************************
```

```
library(tidyr)
library(broman)
```

```
# Order of bodies
ORDER_OF_BODIES = c('Saturn',
                    'Jupiter',
                    'Mars',
                    'Sun',
                    'Venus',
                    'Mercury',
                    'Moon')
names(ORDER_OF_BODIES) <- c("Sa", "J", "Mar", "Su", "V", "Mer", "Moo")
# For plots
theme_basic <- function () {
  theme_bw(base_size=12) %+replace%
    theme(
      axis.text=element_text(colour="black")
    ) %+replace%
    theme(
      panel.grid=element_blank()
```

```
    )
}
```

Singles

We read in the sentiments of the singles.

```
singles.sentiments <- read.csv('../../data/singles-qualities.csv',
                               header=T,
                               stringsAsFactors = F)
# Get in useful format
singles.sentiments.df <- singles.sentiments %>% group_by(PLANET, SENTIMENT) %>%
  summarise(count=n()) %>%
  mutate(total=sum(count),
         prop=count/total)
# remove 'total' variable
singles.sentiments.df$total <- NULL
# Make sure list is complete (some sentiments are missing where none listed in input data)
full_data <- expand.grid(PLANET=singles.sentiments.df$PLANET, SENTIMENT=singles.sentiments.df$SENTIMENT
singles.sentiments.df <- unique(left_join(tbl_df(full_data),singles.sentiments.df))
```

```
## Joining, by = c("PLANET", "SENTIMENT")
```

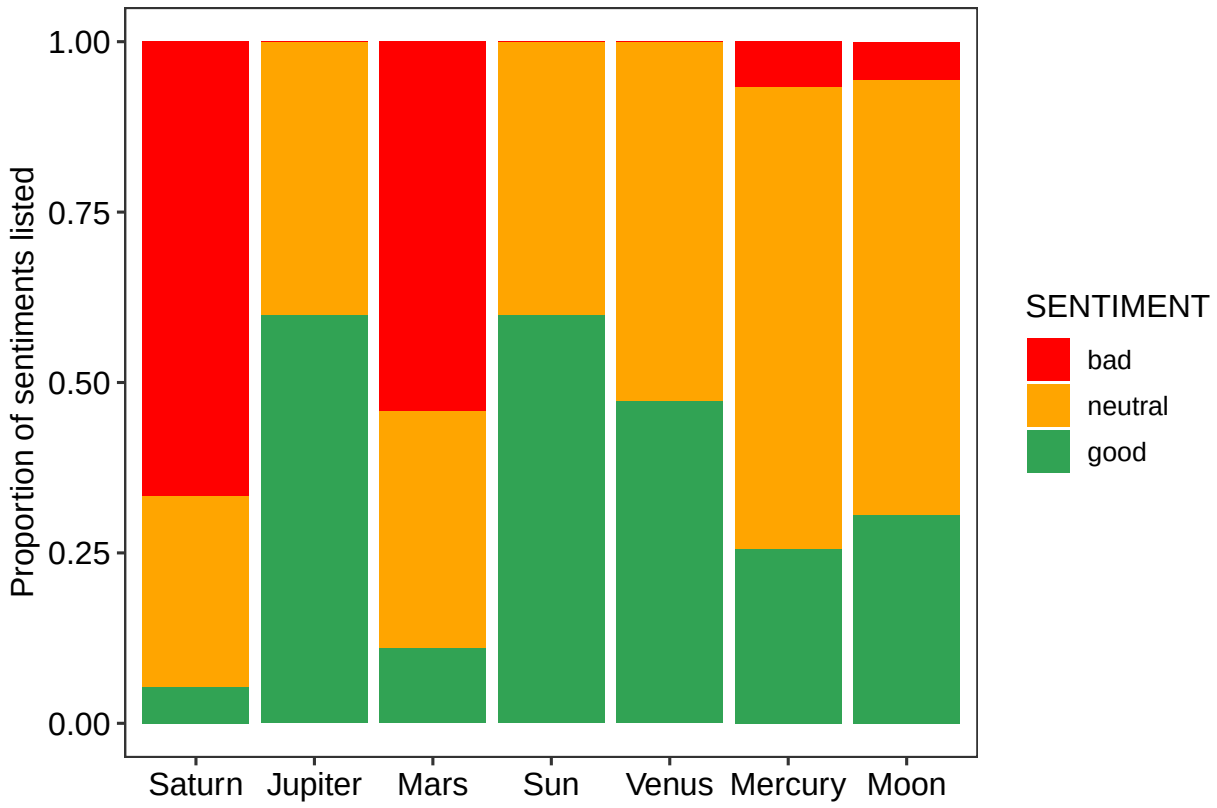
```
## Warning: Column `PLANET` joining factor and character vector, coercing into
## character vector
```

```
## Warning: Column `SENTIMENT` joining factor and character vector, coercing into
## character vector
```

```
singles.sentiments.df$count[is.na(singles.sentiments.df$count)] <- 0
singles.sentiments.df$prop[is.na(singles.sentiments.df$prop)] <- 0
# Order planets and sentiments
singles.sentiments.df$PLANET <- ordered(singles.sentiments.df$PLANET,
                                        levels=ORDER_OF_BODIES)
singles.sentiments.df$SENTIMENT <- ordered(singles.sentiments.df$SENTIMENT,
                                         levels=c("bad", "neutral", "good"))
# Overall sentiment index
singles.overall.sentiments <- singles.sentiments.df %>% group_by(PLANET) %>%
  summarise(sentiment=prop[which(SENTIMENT=="good")]-prop[which(SENTIMENT=="bad")])
```

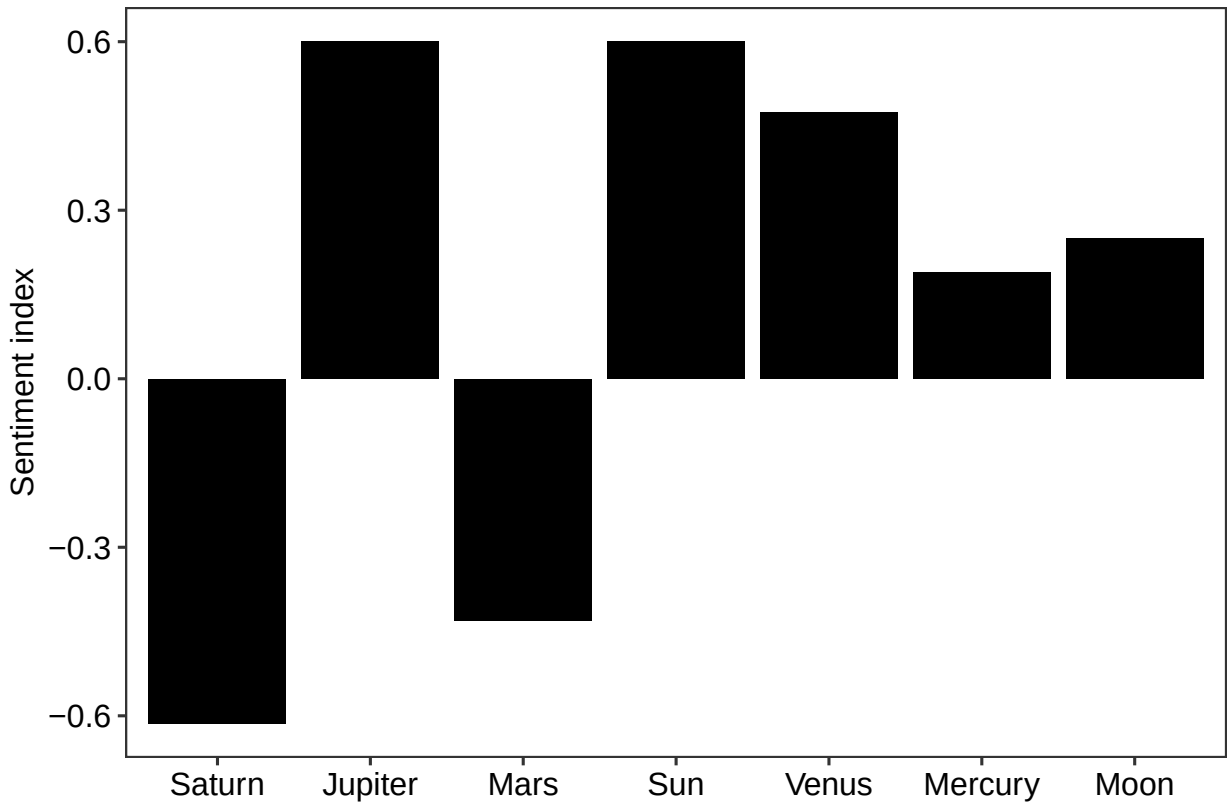
We then plot the sentiments of the singles.

```
# Plot them
ggplot(singles.sentiments.df, aes(PLANET, prop, fill=SENTIMENT))+
  geom_bar(stat="identity")+
  theme_basic()+
  xlab("")+
  scale_fill_manual(values=c("red", "orange", "#31a354"))+
  ylab("Proportion of sentiments listed")
```



Also plot overall sentiment index (good - bad).

```
ggplot(singles.overall.sentiments, aes(PLANET, sentiment))+
  geom_bar(stat="identity", fill="black")+
  theme_basic()+
  xlab("")+
  ylab("Sentiment index")
```



Doubles

We read in the sentiments of the double conjunctions (e.g. Saturn and Jupiter).

```
doubles.sentiments <- read.csv('../../data/doubles-qualities.csv',
                               header=T,
                               stringsAsFactors = F)
# Get in useful format
doubles.sentiments.df <- doubles.sentiments %>% group_by(DOUBLE, SENTIMENT) %>%
  summarise(count=n()) %>%
  mutate(total=sum(count),
         prop=count/total)
# remove 'total' variable
doubles.sentiments.df$total <- NULL
# Make sure list is complete (some sentiments are missing where none listed in input data)
full_data <- expand.grid(DOUBLE=doubles.sentiments.df$DOUBLE, SENTIMENT=doubles.sentiments.df$SENTIMENT
doubles.sentiments.df <- unique(left_join(tbl_df(full_data),doubles.sentiments.df))
```

```
## Joining, by = c("DOUBLE", "SENTIMENT")
```

```
## Warning: Column `DOUBLE` joining factor and character vector, coercing into
## character vector
```

```
## Warning: Column `SENTIMENT` joining factor and character vector, coercing into
## character vector
```

```
doubles.sentiments.df$count[is.na(doubles.sentiments.df$count)] <- 0
doubles.sentiments.df$prop[is.na(doubles.sentiments.df$prop)] <- 0


# Order planets and sentiments
```

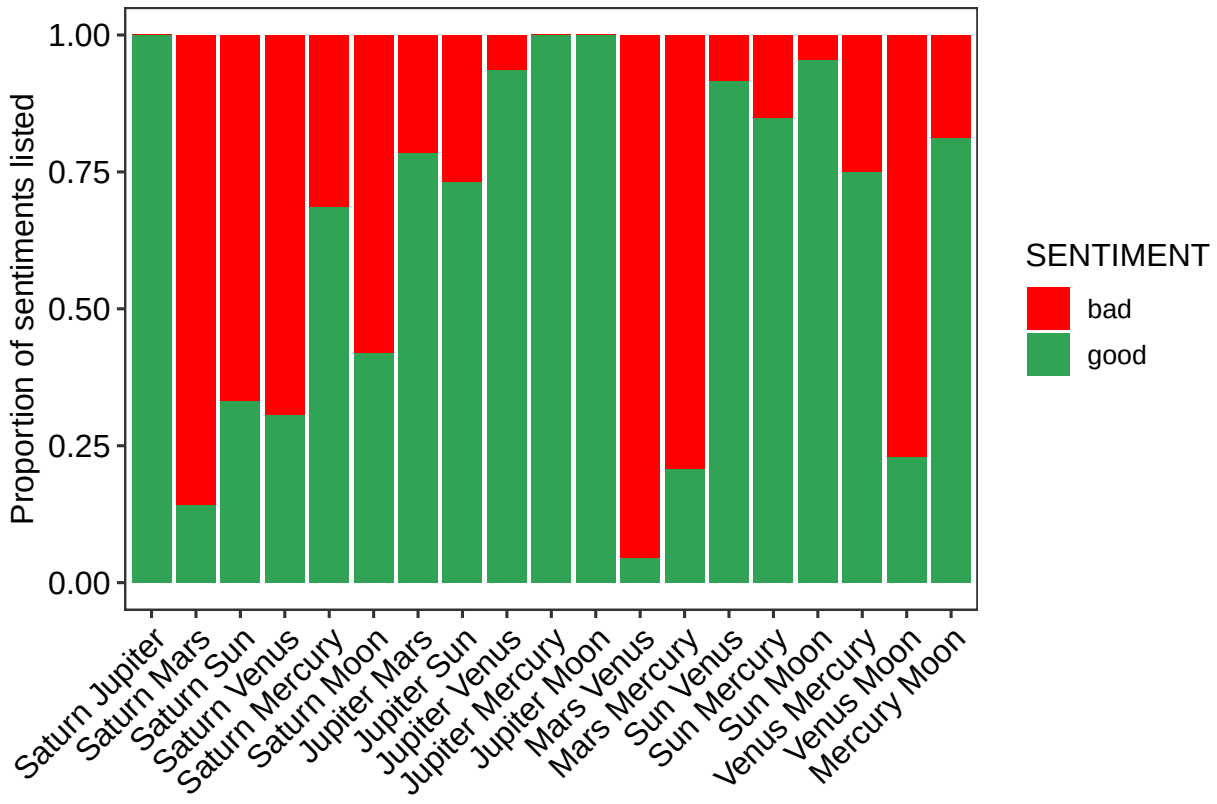
```
doubles.sentiments.df$body.1 <- sapply(stringr::str_split(doubles.sentiments.df$DOUBLE, pattern=" "),
                                       function(x) x[1])
doubles.sentiments.df$body.1 <- ordered(ORDER_OF_BODIES[doubles.sentiments.df$body.1],
                                        levels=ORDER_OF_BODIES)
doubles.sentiments.df$body.2 <- sapply(stringr::str_split(doubles.sentiments.df$DOUBLE, pattern=" "),
                                       function(x) x[2])
doubles.sentiments.df$body.2 <- ordered(ORDER_OF_BODIES[doubles.sentiments.df$body.2],
                                        levels=ORDER_OF_BODIES)
doubles.sentiments.df$bodies.sorted <- sapply(1:nrow(doubles.sentiments.df),
                                              function(x)
                                                paste(as.character(sort(unlist(c(doubles.sentiments.df[x,"body.1"
                                                                                 doubles.sentiments.df[x,"body.2"
                                                      collapse=" "))
# Make body 1 and 2 be in order of bodies as expected
doubles.sentiments.df$body.1 <- gsub(" .*", "", doubles.sentiments.df$bodies.sorted)
doubles.sentiments.df$body.2 <- gsub(".* ", "", doubles.sentiments.df$bodies.sorted)
doubles.sentiments.df$body.1 <- ordered(doubles.sentiments.df$body.1,
                                        levels=ORDER_OF_BODIES)
doubles.sentiments.df$body.2 <- ordered(doubles.sentiments.df$body.2,
                                        levels=ORDER_OF_BODIES)
doubles.sentiments.df$order.body.string <- paste0(as.numeric(doubles.sentiments.df$body.1),
                                                  as.numeric(doubles.sentiments.df$body.2))
# Order bodies again (this is hacky but it works)
doubles.sentiments.df$bodies.sorted <- ordered(doubles.sentiments.df$bodies.sorted,
                                               levels=unique(doubles.sentiments.df$bodies.sorted[order(do

doubles.sentiments.df$SENTIMENT <- ordered(doubles.sentiments.df$SENTIMENT,
                                           levels=c("bad", "good"))

# Also make a dataframe with an overall sentiment index
doubles.overall.sentiments <- doubles.sentiments.df %>% group_by(body.1, body.2, bodies.sorted, order.bo
  summarise(sentiment=prop[which(SENTIMENT=="good")]-prop[which(SENTIMENT=="bad")])
```

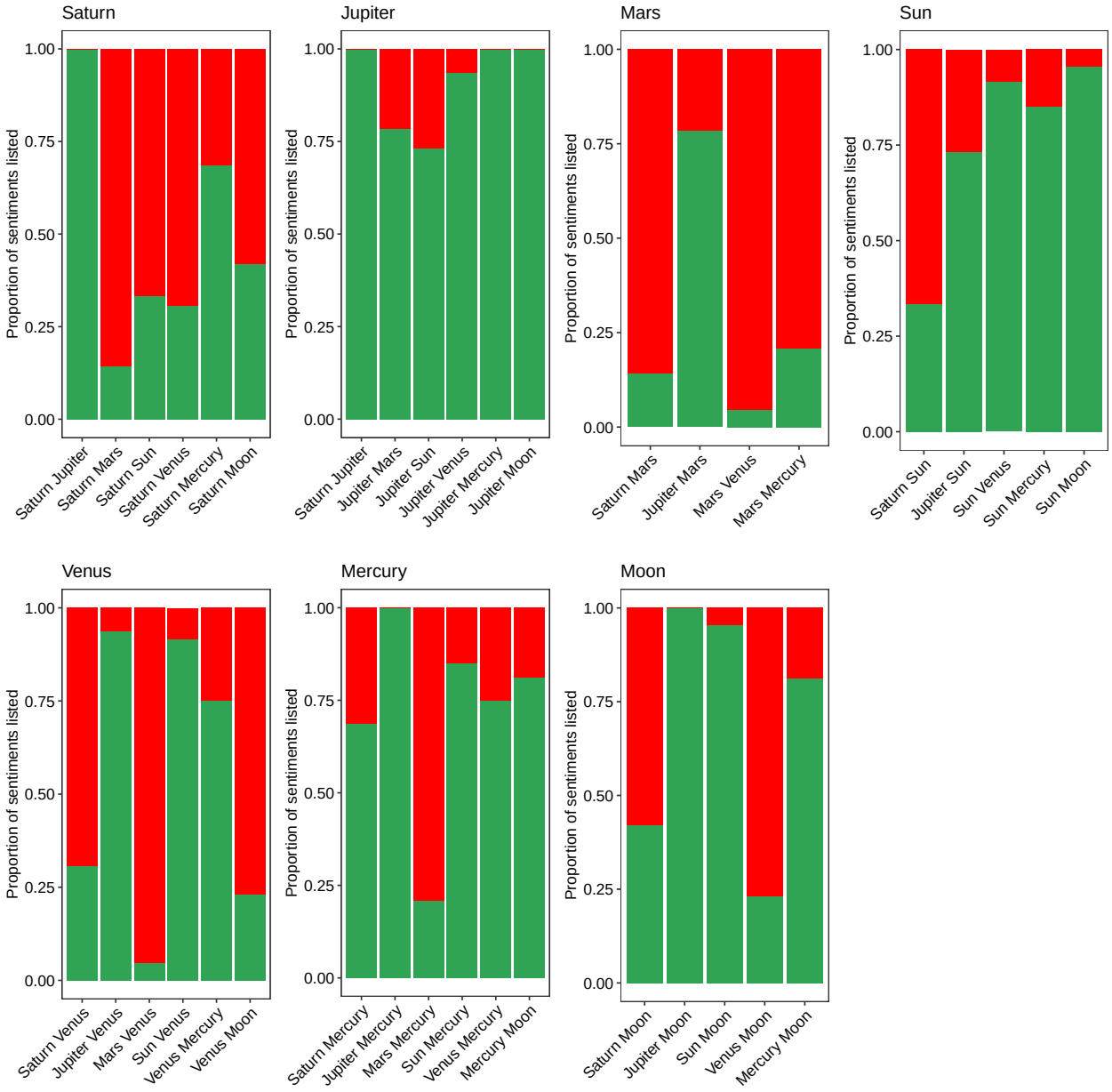
We then plot the sentiments of the doubles.

```
# Plot them
ggplot(doubles.sentiments.df, aes(bodies.sorted, prop, fill=SENTIMENT))+
  geom_bar(stat="identity")+
  theme_basic()+
  xlab("")+
  scale_fill_manual(values=c("red",  "#31a354"))+
  ylab("Proportion of sentiments listed")+
  theme(axis.text.x=element_text(angle=45, hjust=1))
```



A function to plot all the doubles involving a particular planet.

```
plotDoublesWithPlanet <- function(planet){
  # Subset to only doubles involving planet
  local.planet.df <- doubles.sentiments.df[grep(planet, doubles.sentiments.df$bodies.sorted),]
  ggplot(local.planet.df, aes(bodies.sorted, prop, fill=SENTIMENT))+
  geom_bar(stat="identity")+
  theme_basic()+
  xlab("")+
  scale_fill_manual(values=c("red", "#31a354"))+
  ylab("Proportion of sentiments listed")+
  theme(axis.text.x=element_text(angle=45, hjust=1))+
    ggtitle(planet)+
    theme(legend.position = "none")
}
# Make all these plots and combine them
p.doubles.saturn <- plotDoublesWithPlanet("Saturn")
p.doubles.jupiter <- plotDoublesWithPlanet("Jupiter")
p.doubles.mars <- plotDoublesWithPlanet("Mars")
p.doubles.sun <- plotDoublesWithPlanet("Sun")
p.doubles.venus <- plotDoublesWithPlanet("Venus")
p.doubles.mercury <- plotDoublesWithPlanet("Mercury")
p.doubles.moon <- plotDoublesWithPlanet("Moon")
cowplot::plot_grid(p.doubles.saturn, p.doubles.jupiter,
                   p.doubles.mars, p.doubles.sun,
                   p.doubles.venus, p.doubles.mercury,
                   p.doubles.moon, nrow=2)
```



Note that Valens misses out two doubles which involve Mars: Mars + Sun, and Mars + Moon.

Predicting doubles from singles

If we know the sentiment proportions associated with single planets, can we predict the sentiment of the doubles?

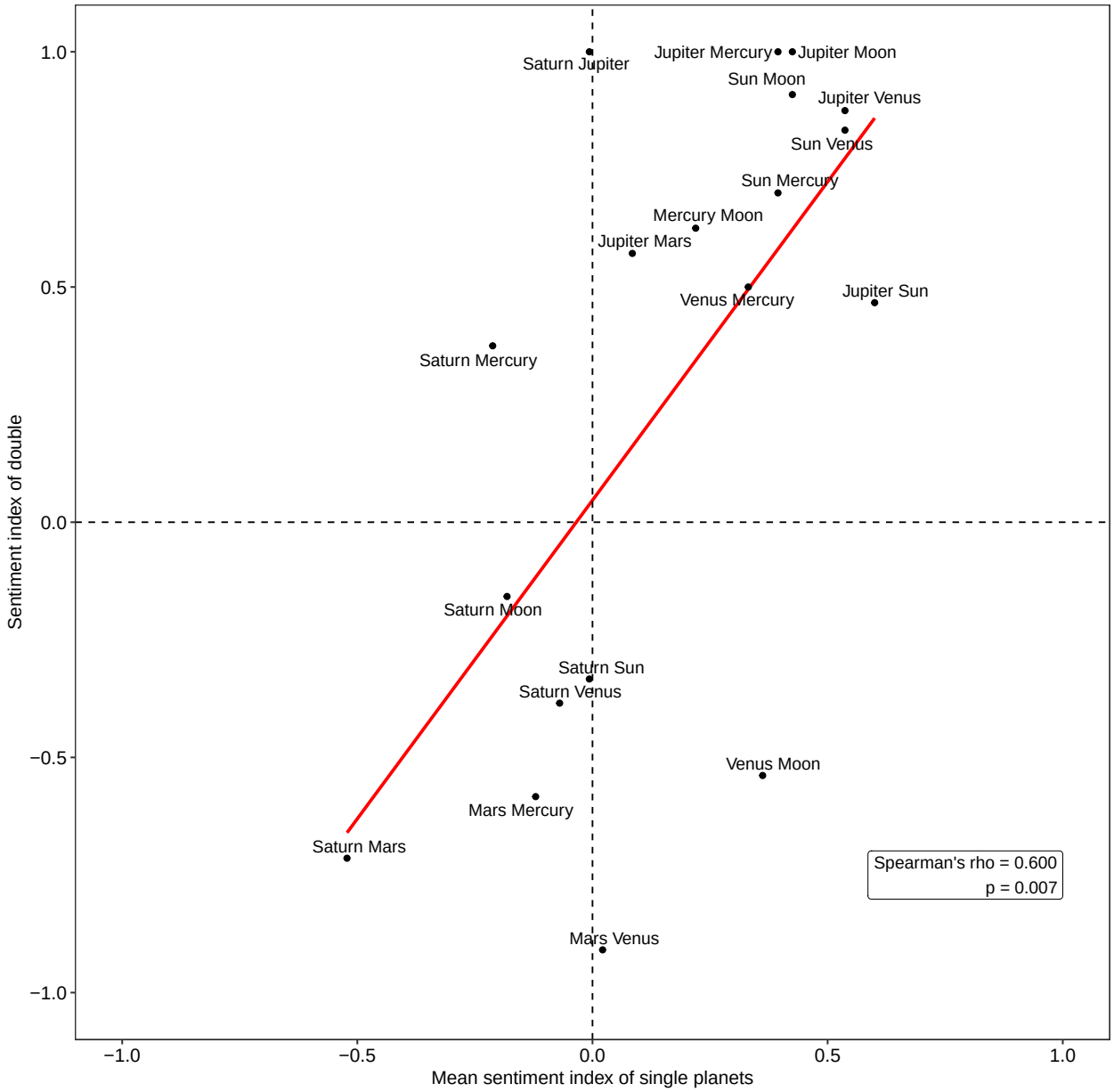
```
# getSingleScores <- function(double, sentiment="good", mean=TRUE, singles=singles.sentiments.df){
#     double.bodies <- ordered(unlist(stringr::str_split(double, pattern=" ")),
#                              levels=ORDER_OF_BODIES)
#     prop <- 0
#     for (body in double.bodies){
#       new.prop <- as.numeric(singles[which(singles$PLANET==body &
#                                      singles$SENTIMENT==sentiment), "prop"]  )
#       prop <- prop +  new.prop
```

```
#     }
#     if (mean==FALSE){
#       return(prop)
#     }
#     else{
#       return(prop/2)
#     }
# }
getSingleScores <- function(double, mean=TRUE, singles=singles.overall.sentiments){
    double.bodies <- ordered(unlist(stringr::str_split(double, pattern=" ")),
                             levels=ORDER_OF_BODIES)
    prop <- 0
    for (body in double.bodies){
      new.prop <- as.numeric(singles[which(singles$PLANET==body), "sentiment"])
      prop <- prop +  new.prop
    }
    if (mean==FALSE){
      return(prop)
    }
    else{
      return(prop/2)
    }
}

doubles.overall.sentiments$mean.single.sentiments <- sapply(doubles.overall.sentiments$bodies.sorted,
                                                            function(x) getSingleScores(x, mean=TRUE))
# Add correlation score
spearman.cor <- cor.test(doubles.overall.sentiments$mean.single.sentiments, doubles.overall.sentiments$
```

```
## Warning in cor.test.default(doubles.overall.sentiments$mean.single.sentiments, :
## Cannot compute exact p-value with ties
```

```
p.sentiments.doubles.singles <- ggplot(doubles.overall.sentiments, aes(mean.single.sentiments, sentimen
    stat_smooth(method="lm", se=FALSE, colour="red")+
  geom_point(colour="black")+
  theme_basic()+
  ggrepel::geom_text_repel(aes(label=bodies.sorted))+
    coord_fixed()+
  xlim(c(-1, 1))+
  ylim(c(-1,1))+
  geom_hline(yintercept = 0, linetype='dashed')+
    geom_vline(xintercept = 0, linetype='dashed')+
  xlab("Mean sentiment index of single planets")+
  ylab("Sentiment index of double")+
  annotate(geom="label", label=paste0("Spearman's rho = ", myround(spearman.cor$estimate, 3), "\n",
                                      "p = ", myround(spearman.cor$p.value, 3)), x=1, y=-0.75, hjust=1)
p.sentiments.doubles.singles
```



We see that there is a strong correlation between the mean sentiment index for the component singles and the sentiment index of the double conjunction.