

Problem Set 1 Code

2025-11-13

Task 1

In the following a descriptive analysis of the “detailed_fish_market_data”, regarding Whiting is conducted. At first the required packages are loaded.

```
## Load packages and dataset
library(readr)
library(dplyr)
library(ggplot2)
library(tidyr)

detailed <- read_tsv("../data/detailed_fish_market_data.txt")
```

Data preparation

Bla Bla Datensatz ec.

```
# delete those rows that have NA for
# "price", "quan", "totr", "tots" and
# filter for whiting (no king)
detailed_whiting <- detailed %>%
  filter(!is.na(pric),
         !is.na(quan),
         !is.na(totr),
         type == "w") %>%
  arrange(date)
```

The next code snippet is not shown in the pdf as those are only technical changes. As the dataset has some observations, that are “NA” or relatively obvious outliers, those are removed first. Especially the two observations, that seem to stem from another dealer are removed.

Dataset for the analysis on daily level is created.

```
# dataset for the daily-level
detailed_whiting_daily <- detailed_whiting %>%
  group_by(date) %>%
  summarise(
    avg_princ = mean(pric),
    totr = first(totr),
    tots = first(tots),
    dayw = first(dayw),
    n_trsact = n(),
    strate = first(tots)/first(totr)
  )
```

Descriptive analysis

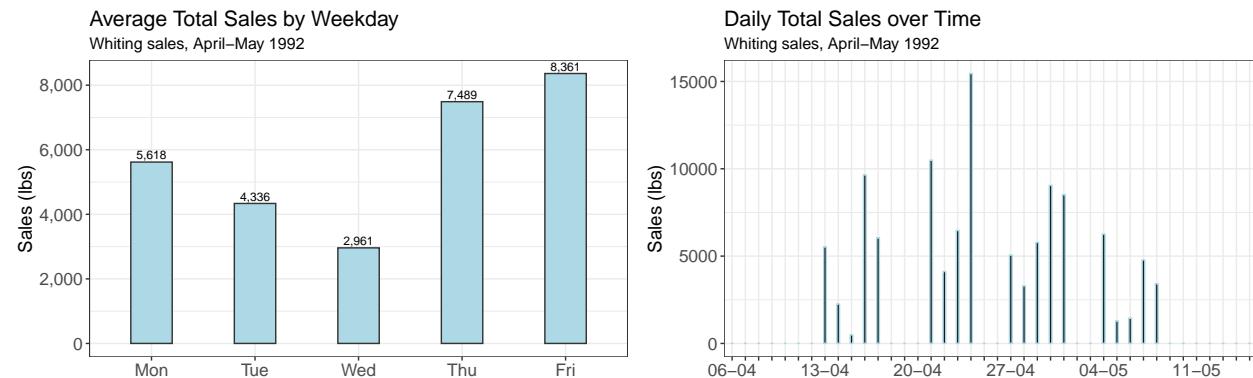
```
####  
# summary of the daily dataset  
####  
detailed_whiting_daily %>%  
  select(totr, tots, n_trsact) %>%  
  summary()
```

```
##          totr          tots        n_trsact  
##  Min.   : 200   Min.   : 490   Min.   : 4.00  
##  1st Qu.: 1990  1st Qu.: 3360  1st Qu.:18.00  
##  Median : 6080  Median : 5535  Median :25.00  
##  Mean   : 5881  Mean   : 5760  Mean   :25.05  
##  3rd Qu.: 7927  3rd Qu.: 7495  3rd Qu.:32.50  
##  Max.   :15940  Max.   :15455  Max.   :57.00
```

As the summary indicates, total sales and therefore the total received amount of Whiting in lbs inherit a large amount of variation. The amount of transaction per day also shows a broad variety of values, with the minimum of four and a maximum of 57.

To gain a first insight in the properties of the total sales of Whiting in the period of April to May 1992 a bar-chart and a time-series plot are used.

```
plot_average_sales_by_weekday  
plot_daily_sales_over_time
```

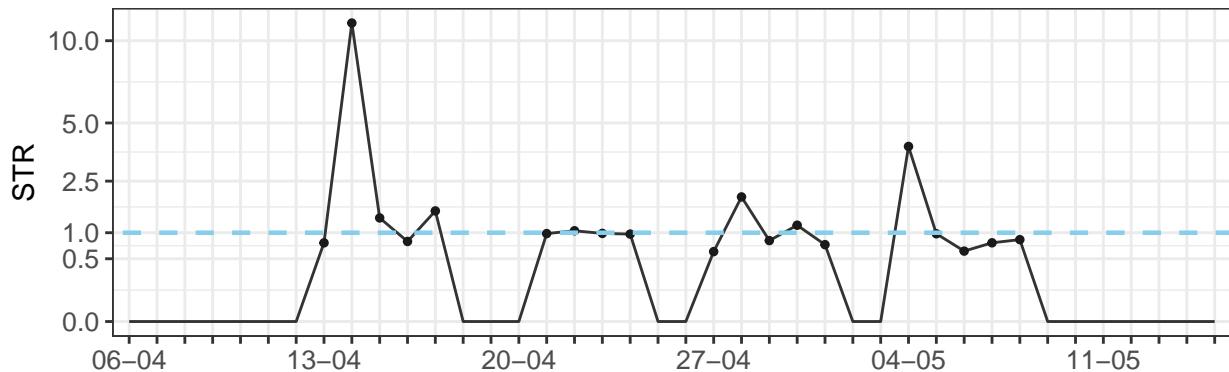


Plot Beschreibung.

```
plot_daily_str_over_time
```

Daily Sell-Through Rate over Time

Whiting sales, April–May 1992



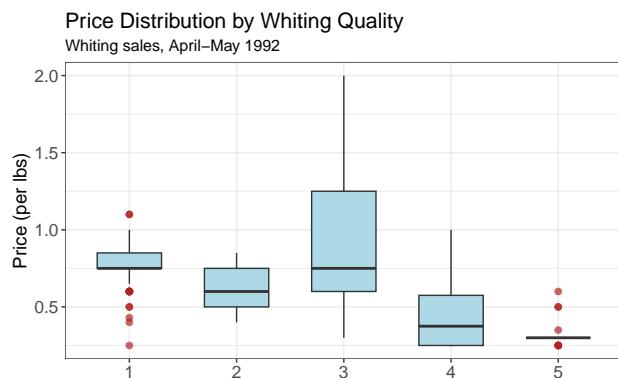
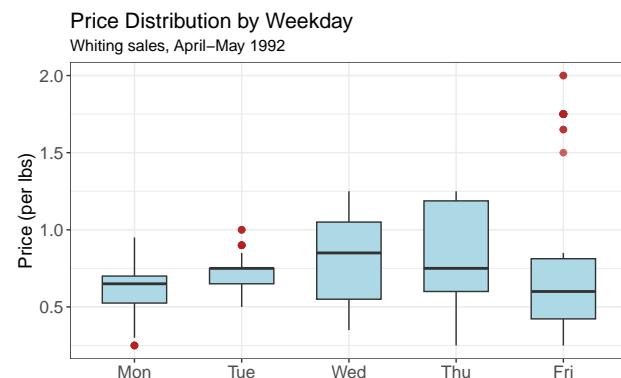
Plot Beschreibung.

```
####  
# correlation between tots (total daily sales) and avg_pric (average price)  
####  
cor(detailed_whiting_daily$avg_pric,detailed_whiting_daily$tots)
```

```
## [1] -0.4236805
```

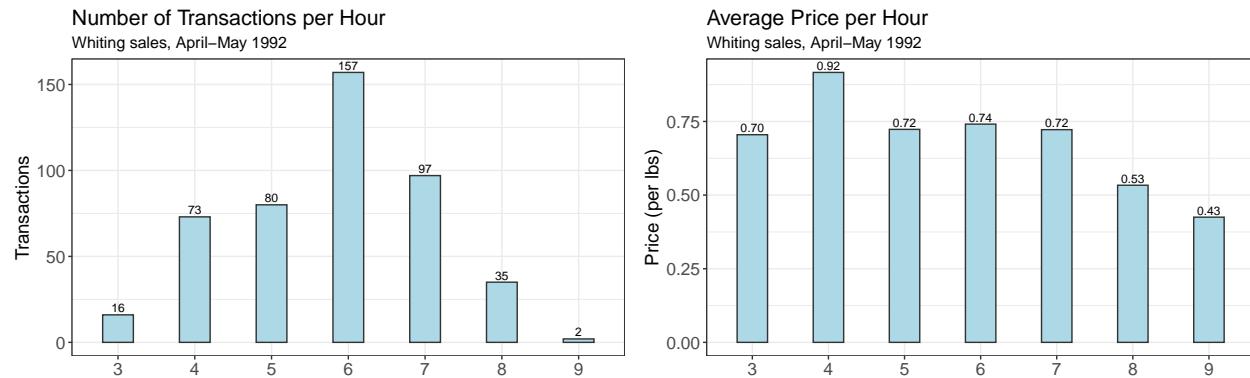
Korrelation Beschreibung.

```
plot_price_distr_by_weekday  
plot_price_distr_by_quality
```



Plot Beschreibung.

```
plot_transactions_per_hour  
plot_average_price_per_hour
```



Plot Beschreibung.