Conjoint Analysis: What is the best laptop design?

Savaş Dayanık

3/9/2021

Contents

Preference of one subject

```
train_hande_df <- read_csv("laptops-hande.csv") %>%
  mutate_if(is.character, factor)
```

```
## Rows: 16 Columns: 9
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (7): Brand, Processor, Speed, Memory, HardDisk, Screen, Price
## dbl (2): Alternative, RANK
```

```
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
test_hande_df <- read_csv("laptops-hepsiburada-hande.csv", comment = "#") %>%
  mutate_if(is.character, factor)
```

```
## Rows: 6 Columns: 9
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (7): Brand, Processor, Speed, Memory, HardDisk, Screen, Price
## dbl (2): Alternative, RANK
```

```
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
summary(train_hande_df)
```

```
##   Alternative        Brand           Processor     Speed    Memory    HardDisk
##  Min.   : 1.00   Casper :4   Intel Core i5:8   1.8 GHz:4   4 GB:8   500 GB:8
##  1st Qu.: 4.75   HP     :4   Intel Core i7:8   2.2 GHz:4   8 GB:8   750 GB:8
##  Median : 8.50   Sony   :4                     2.4 GHz:4
##  Mean   : 8.50   Toshiba:4                     2.6 GHz:4
##  3rd Qu.:12.25
##  Max.   :16.00
##      Screen      Price        RANK
##  15.6 in:8   1500 TL:4   Min.   : 1.00
##  17.3 in:8   1700 TL:4   1st Qu.: 4.75
##              2000 TL:4   Median : 8.50
##              3500 TL:4   Mean   : 8.50
##                          3rd Qu.:12.25
##                          Max.   :16.00
```

```
train_hande_df %>%
  arrange(RANK)
```

```
## # A tibble: 16 x 9
##    Alternative Brand   Processor     Speed   Memory HardDisk Screen  Price  RANK
##          <dbl> <fct>   <fct>         <fct>   <fct>  <fct>    <fct>   <fct> <dbl>
##  1           2 Sony    Intel Core i5 2.4 GHz 8 GB   750 GB   17.3 in 1500~     1
##  2          12 Toshiba Intel Core i5 2.2 GHz 8 GB   750 GB   15.6 in 1700~     2
##  3           8 HP      Intel Core i5 2.6 GHz 8 GB   750 GB   15.6 in 2000~     3
##  4          16 HP      Intel Core i7 2.4 GHz 8 GB   500 GB   17.3 in 1700~     4
##  5          14 Casper  Intel Core i7 2.2 GHz 8 GB   500 GB   15.6 in 1500~     5
##  6           3 Casper  Intel Core i5 2.6 GHz 4 GB   500 GB   17.3 in 1700~     6
##  7          13 Toshiba Intel Core i7 1.8 GHz 8 GB   500 GB   17.3 in 2000~     7
##  8           1 HP      Intel Core i5 1.8 GHz 4 GB   500 GB   15.6 in 1500~     8
##  9          10 Toshiba Intel Core i7 2.6 GHz 4 GB   750 GB   17.3 in 1500~     9
## 10           5 Sony    Intel Core i5 2.2 GHz 4 GB   500 GB   17.3 in 2000~    10
## 11           6 Sony    Intel Core i7 1.8 GHz 4 GB   750 GB   15.6 in 1700~    11
## 12           7 Casper  Intel Core i7 2.4 GHz 4 GB   750 GB   15.6 in 2000~    12
## 13           4 Sony    Intel Core i7 2.6 GHz 8 GB   500 GB   15.6 in 3500~    13
## 14           9 HP      Intel Core i7 2.2 GHz 4 GB   750 GB   17.3 in 3500~    14
## 15          15 Toshiba Intel Core i5 2.4 GHz 4 GB   500 GB   15.6 in 3500~    15
## 16          11 Casper  Intel Core i5 1.8 GHz 8 GB   750 GB   17.3 in 3500~    16
```

```
options("contrasts")
```

```
## $contrasts
##         unordered           ordered
## "contr.treatment"      "contr.poly"
```

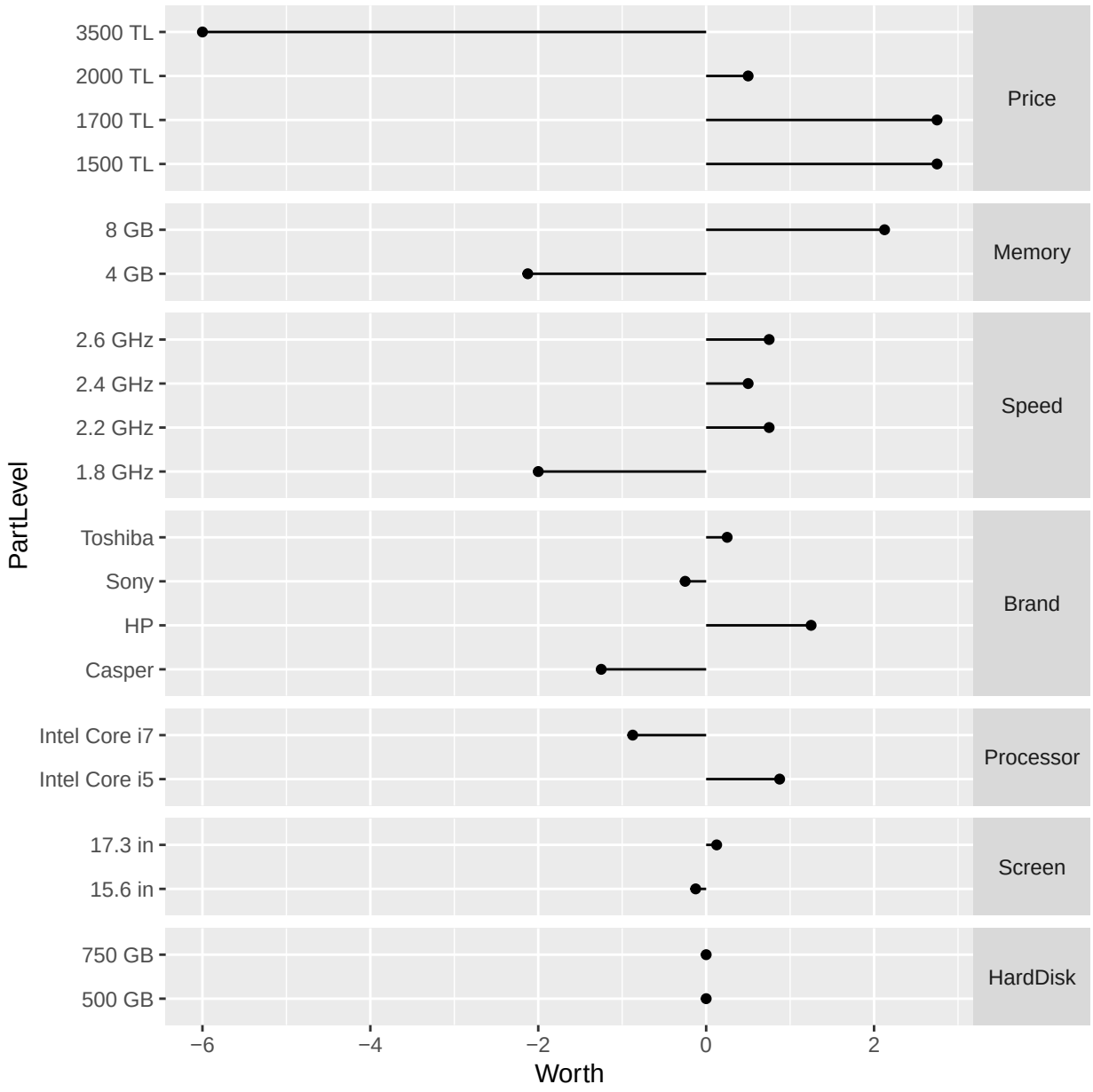
```
options(contrasts = c("contr.sum", "contr.poly"))
```

Estimate Hande's part worth and importance scores

```
coef_df <- lm(-RANK ~ ., select(train_hande_df, - Alternative)) %>%
  summary() %>%
  coef() %>%
  as_tibble(rownames = "Part") %>%
  select(Part, Estimate) %>%
  filter(Part != "(Intercept)") %>%
  rename(Worth = Estimate)

parts_df <- coef_df %>%
  mutate(ID = as.numeric(str_extract(Part, "\\d+")),
         Part = str_extract(Part, "\\D*")) %>%
  nest(data=-Part) %>%
  mutate(data2 = map(data, ~rbind(.x, tibble(Worth = -sum(.x$Worth),
                                             ID = max(.x$ID + 1))))) %>%
  select(-data) %>%
  unnest(data2) %>%
  mutate(PartLevel = map2_chr(Part, ID, ~levels(train_hande_df[[.x]])[.y])) %>%
  group_by(Part) %>%
  mutate(Importance = diff(range(Worth)))
```

Plot Hande's part worth and importance scores

```
parts_df %>%
  ggplot(aes(PartLevel, Worth)) +
  geom_point() +
  geom_linerange(aes(ymin=0, ymax = Worth)) +
  coord_flip() +
  facet_grid(reorder(Part, -Importance) ~ ., space="free", scale="free") +
  theme(strip.text.y = element_text(angle=0))
```



Estimate market participant part worths and and importance scores

Now repeat the same analysis for all subjects who filled up the questionnaire.

```
# d2 <- read_csv("conjoint/all/laptops train survey results-all.csv") %>%
train_df <- read_excel("laptops train survey results-all-coded.xlsx") %>%
  mutate_if(is.character, factor) %>%
  mutate(id = as.character(id),
         Alternative = factor(Alternative, level = levels(Alternative)[order(parse_number(levels(Alterna
         Price = factor(Price, level = levels(Price)[order(parse_number(levels(Price)))]))

train_df %>% summary()
```

```
##       id             Alternative       Brand             Processor
##  Length:2960        1      : 185   Casper :740   Intel Core i5:1480
##  Class :character   2      : 185   HP     :740   Intel Core i7:1480
##  Mode  :character   3      : 185   Sony   :740
##                     4      : 185   Toshiba:740
##                     5      : 185
##                     6      : 185
##                     (Other):1850
##      Speed       Memory         HardDisk        Screen           Price
##  1.8 GHz:740   16 GB:1480   1.000 GB:1480   15.6 in:1480   6.000 TL :740
##  2.2 GHz:740   8 GB :1480   750 GB  :1480   17.3 in:1480   7.000 TL :740
##  2.4 GHz:740                                               8.500 TL :740
##  2.6 GHz:740                                               10.000 TL:740
##
##
##
##       RANK
##  Min.   : 1.000
##  1st Qu.: 4.000
##  Median : 8.000
##  Mean   : 8.491
##  3rd Qu.:12.000
##  Max.   :16.000
##
```

```
test_df <- read_excel("laptops test survey results-all-coded.xlsx") %>%
    mutate_if(is.character, factor)

test_df %>% summary()
```

```
##            id       Alternative     Brand             Processor       Speed
##  y14s208s001:   6   1:185       Casper :370   Intel Core i5:740   1.8 GHz:370
##  y14s209s001:   6   2:185       HP     :185   Intel Core i7:370   2.2 GHz:185
##  y14s209s002:   6   3:185       Sony   :185                       2.4 GHz:370
##  y14s209s003:   6   4:185       Toshiba:370                       2.6 GHz:185
##  y14s209s005:   6   5:185
##  y14s209s006:   6   6:185
##  (Other)    :1074
##    Memory         HardDisk       Screen          Price          RANK
##  16 GB:370   1.000 GB:555   15.6 in:925   10.000 TL:185   Min.   :1.0
##  8 GB :740   750 GB  :555   17.3 in:185   6.000 TL :370   1st Qu.:2.0
##                                           7.000 TL :370   Median :3.5
##                                           8.500 TL :185   Mean   :3.5
##                                                           3rd Qu.:5.0
##                                                           Max.   :6.0
##
```

```
market_df <- train_df %>%
  nest(train_data = -id) %>%
  mutate(
    # learn individual preferences of each respondent / subject
    model = map2(train_data, id,  ~{
      # cat ("ID=", .y, "\n")
      d <- .x
      lm(-RANK ~ ., select(d, - Alternative))
```

```
  }),
cat("model     : ranks are learned with multiple regression.\n"),

# compile the part worths
pworth = pmap(list(model, train_data, id),  ~{
  # cat ("ID=", ..3, "\n")
  d <- ..2
  linmod <- ..1

  coef_df <- linmod %>%
  summary() %>%
  coef() %>%
  as_tibble(rownames = "Part") %>%
  select(Part, Estimate) %>%
  filter(Part != "(Intercept)") %>%
  rename(Worth = Estimate)

  parts_df <- coef_df %>%
    mutate(ID = as.numeric(str_extract(Part, "\\d+")),
           Part = str_extract(Part, "\\D*")) %>%
    nest(data=-Part) %>%
    mutate(data2 = map(data, ~rbind(.x, tibble(Worth = -sum(.x$Worth),
                                               ID = max(.x$ID + 1))))) %>%
    select(-data) %>%
    unnest(data2) %>%
    mutate(PartLevel = map2_chr(Part, ID, ~levels(d[[.x]])[.y])) %>%
    group_by(Part) %>%
    mutate(Importance = diff(range(Worth)),
           # below we pad from left numerical Part levels in order
           # to maintain the numerical order of levels in partworth plots
           # stored in pwplot.
           PartLevel = if (any(str_detect(PartLevel, "\\d"))) {
             str_pad(PartLevel, width = max(str_length(PartLevel)))
           } else {
               PartLevel
             }
           ) %>%
    ungroup()
  parts_df
  }),
cat("pworth    : part worth scores are calculated.\n"),

# extract importance scores
importance = map2(pworth, id, ~{
  # cat ("ID=", .y, "\n")
  .x %>%
    distinct(Part, Importance) %>%
    mutate(Importance = 100*Importance/sum(Importance))
}),
cat("importance: part importance scores are calculated.\n"),

# visualize the part worths
pwplot = map2(pworth, id,  ~{
```

```
      # cat ("ID=", .y, "\n")
      parts_df <- .x

    parts_df %>%
        ggplot(aes(PartLevel, Worth)) +
        geom_point() +
        geom_linerange(aes(ymin=0, ymax = Worth)) +
        coord_flip() +
        facet_grid(reorder(Part, -Importance) ~ ., space="free", scale="free") +
        theme(strip.text.y = element_text(angle=0)) +
        labs(title = .y)

  }),
  cat("pwplot    : part worth plots are ready.\n"),
  ) %>%

  # append test data
  left_join(nest(test_df, test_data = -id), by = "id") %>%
  mutate(
    # rank predictions on test
    test_pred = pmap(list(model, test_data, id), ~{
      # cat(..3, " ")
      pred <- ..1 %>%
        predict(newdata=..2) %>%
        `*` (-1)
      rank <- pred %>%
        rank(ties.method = "first")

      ..2 %>%
        select(Alternative, RANK) %>%
        mutate(PREDraw = pred, PREDrank = rank)
      }),
    cat("test_pred : rankings on test data are predicted.\n")
  )
```

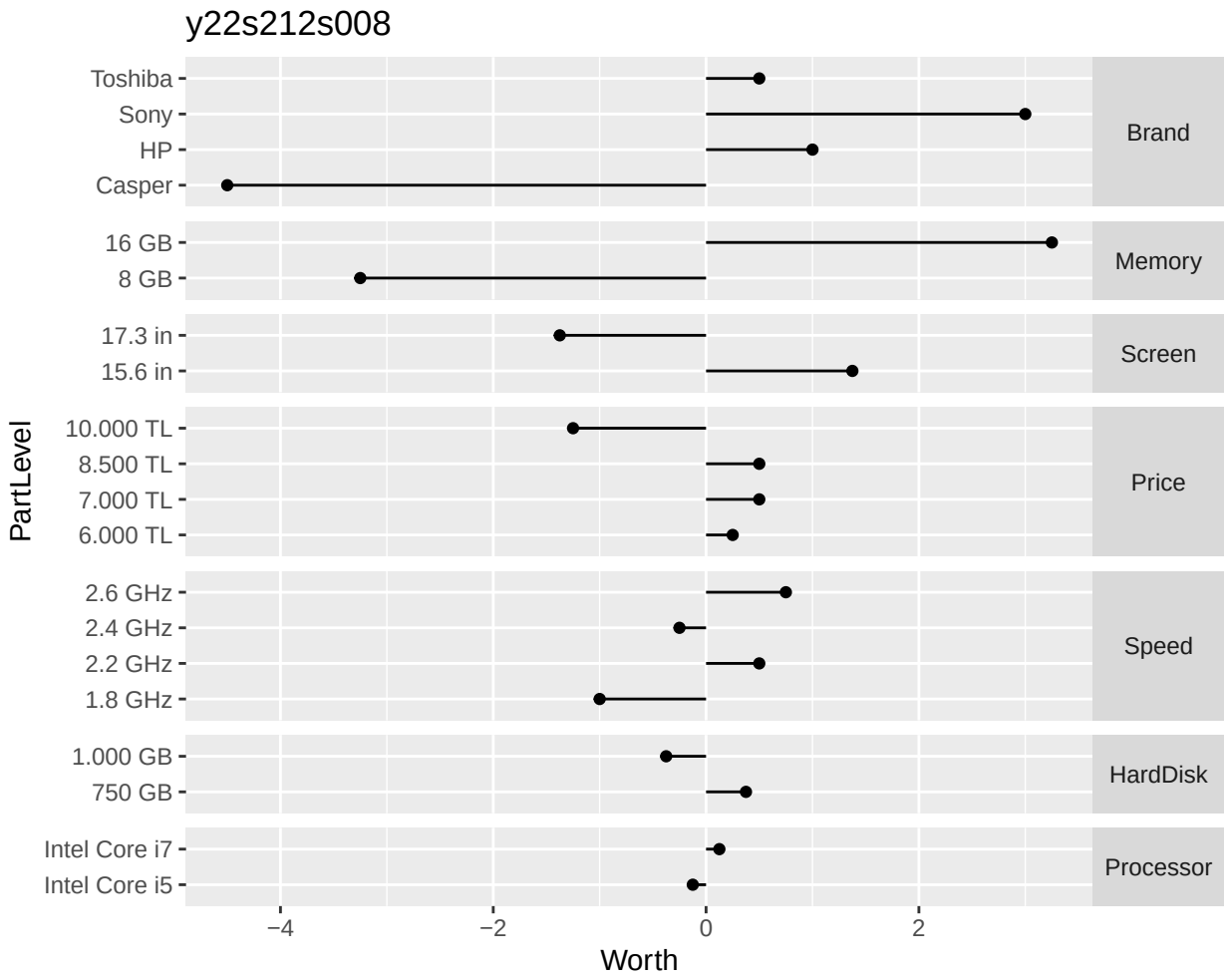
```
## model     : ranks are learned with multiple regression.
## pworth    : part worth scores are calculated.
## importance: part importance scores are calculated.
## pwplot    : part worth plots are ready.
## test_pred : rankings on test data are predicted.
```

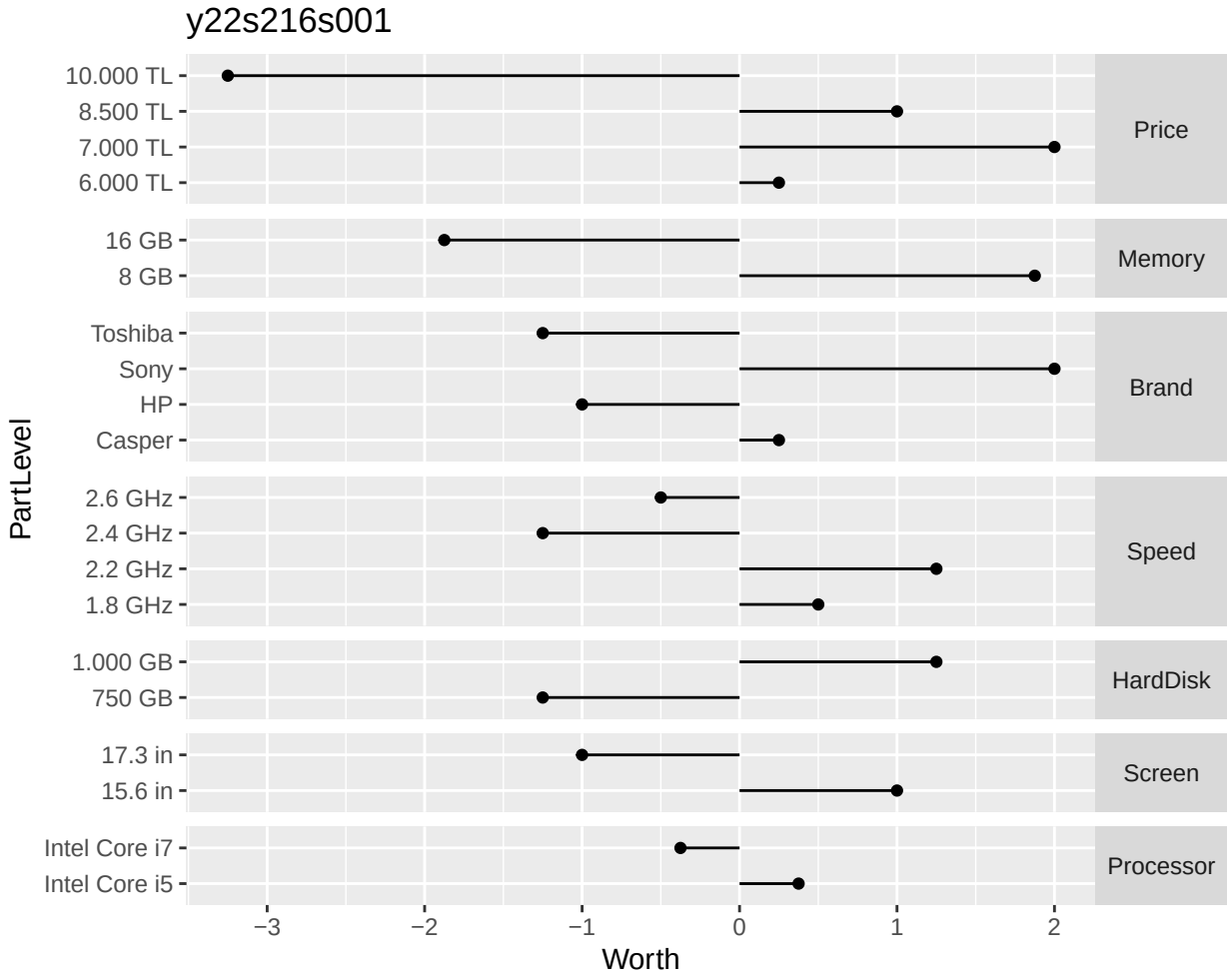
```
  #   test_observed_pred = map2(test_data, test_pred, ~ {
  #     .x %>%
  #       select(Alternative, RANK) %>%
  #       mutate(PRED = .y)
  #   }),
  #   cat("test_opred: observed and predicted rankings are combined.\n"),
  # )
```

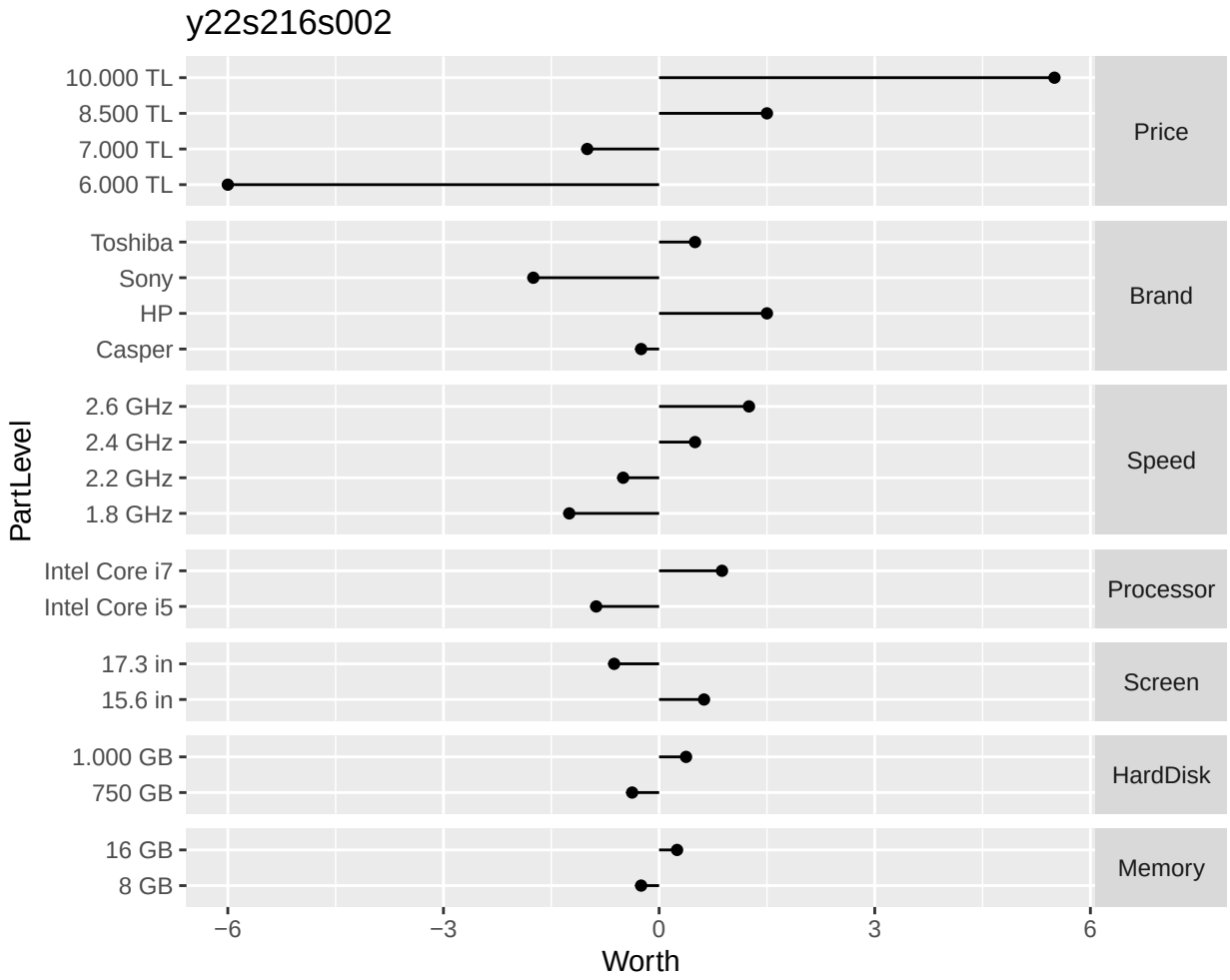
Plot a few market participant part worths and and importance scores

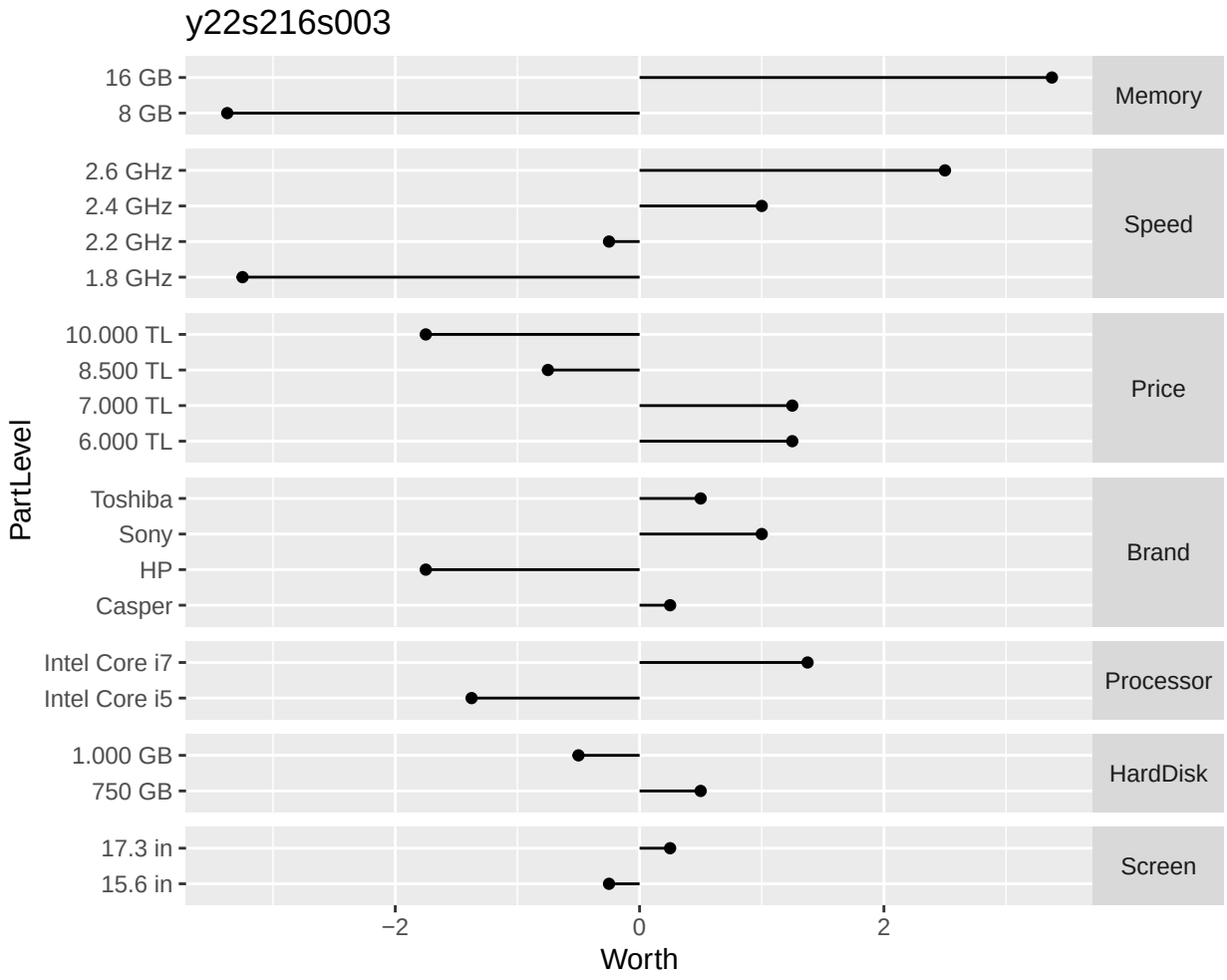
Show a few part worth plots

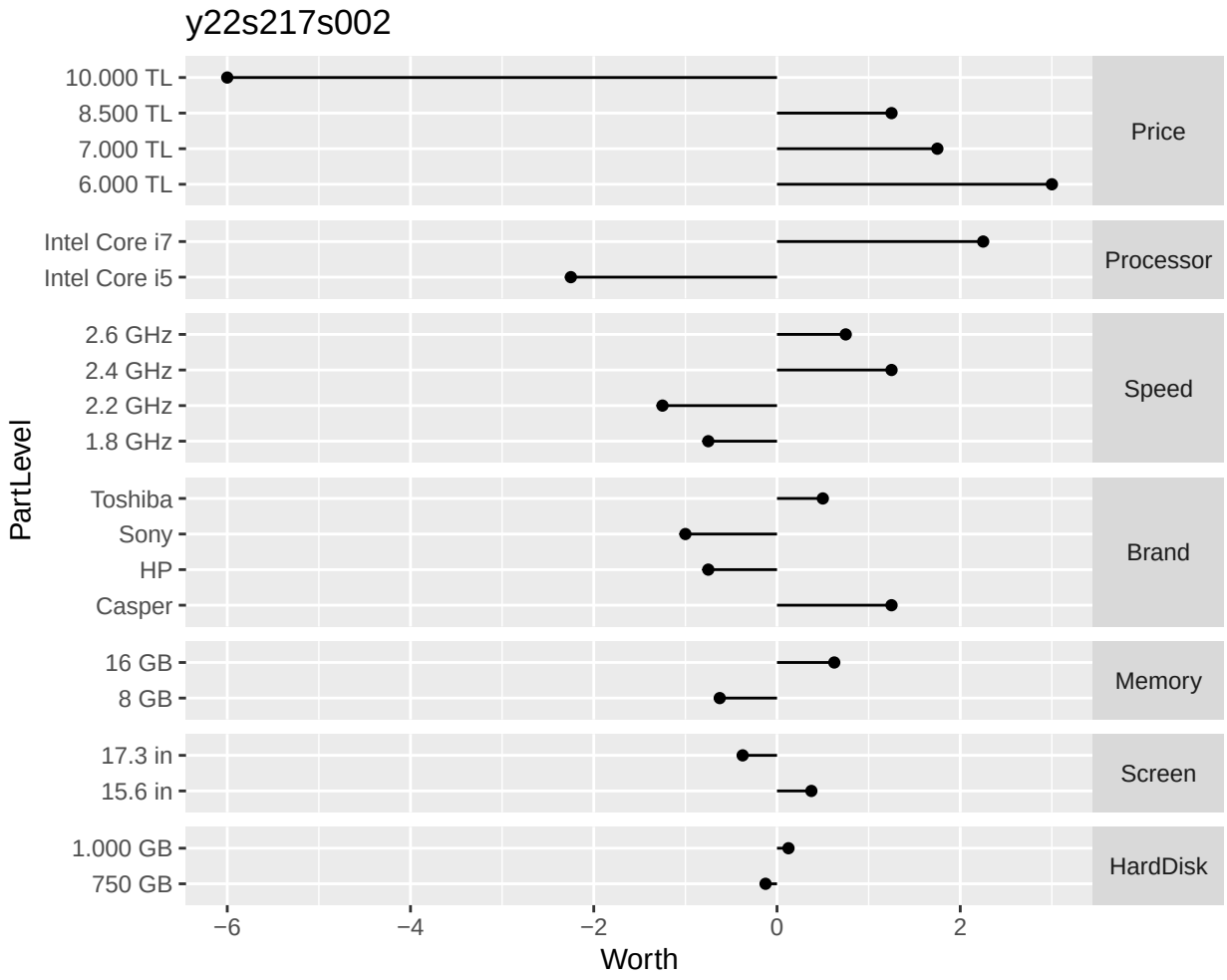
```
# for (i in 1:10) plot(market_df$pwplot[[i]])

market_df %>%
  filter(str_detect(id, "y22")) %>%
  pull(pwplot) %>%
  walk(plot)
```

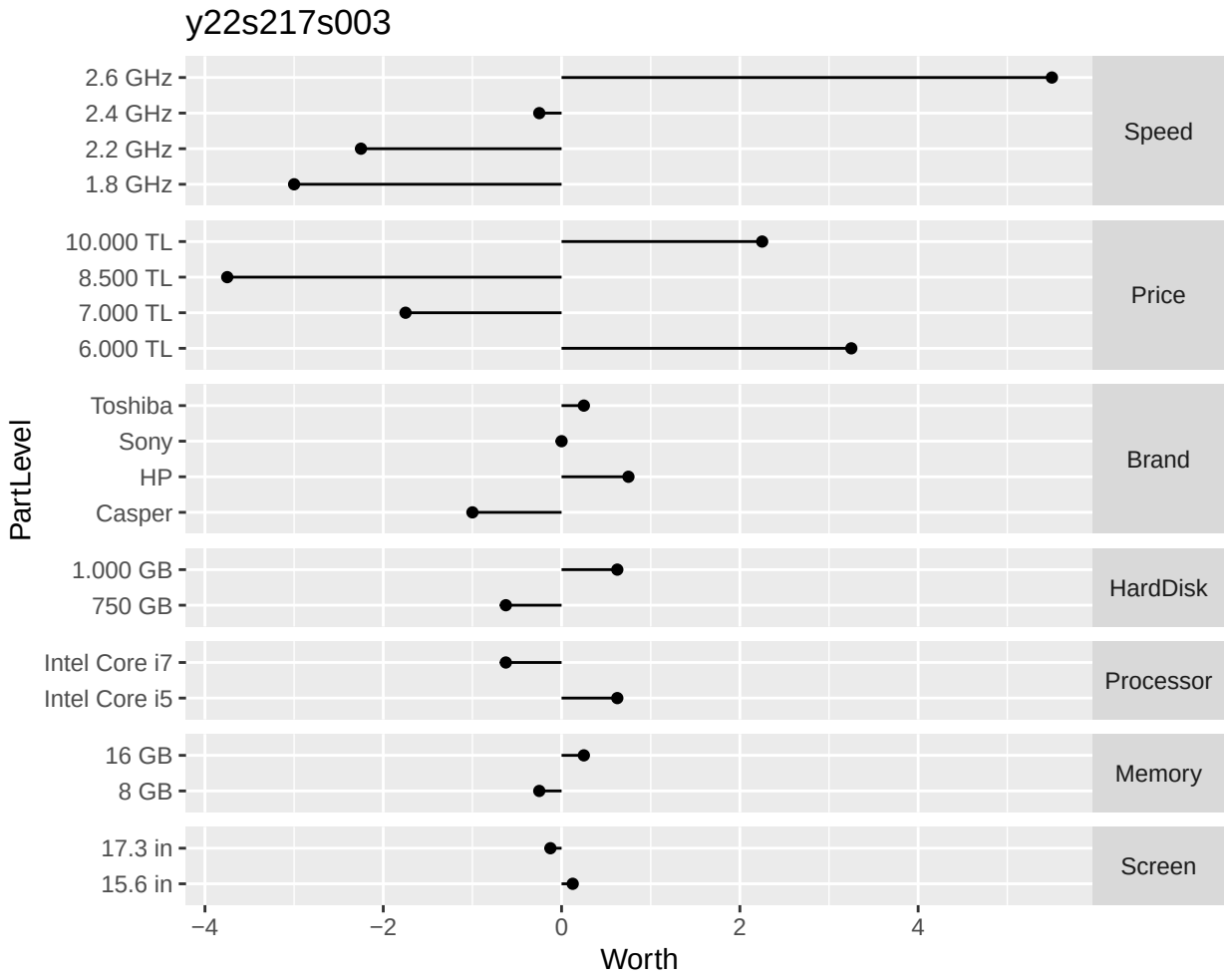


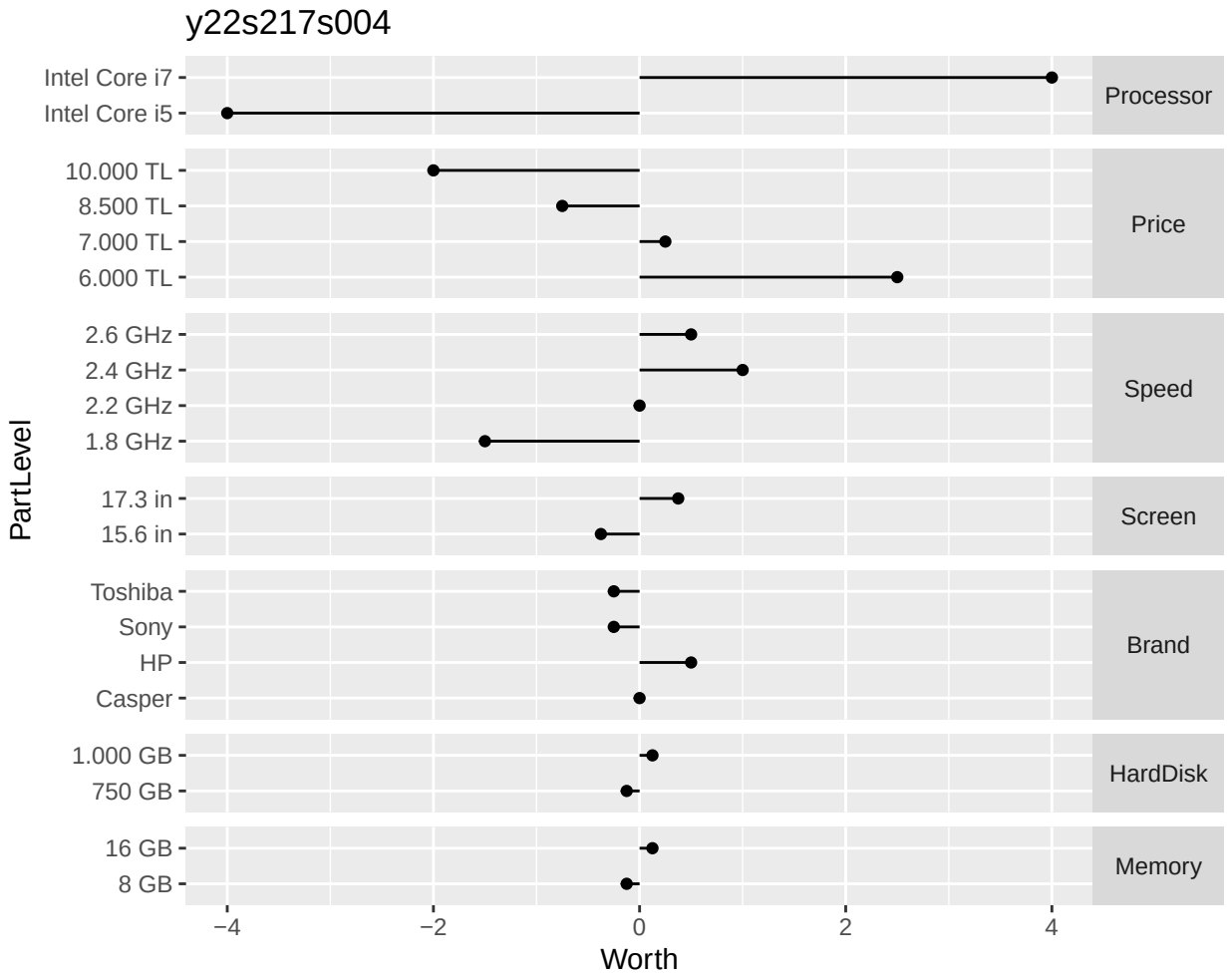

y22s216s001
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
16 GB
8 GB
Memory
Toshiba
Sony
HP
Casper
Brand
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
1.000 GB
750 GB
HardDisk
17.3 in
15.6 in
Screen
Intel Core i7
Intel Core i5
Processor
PartLevel
−3
−2
−1
0
1
2
Worth

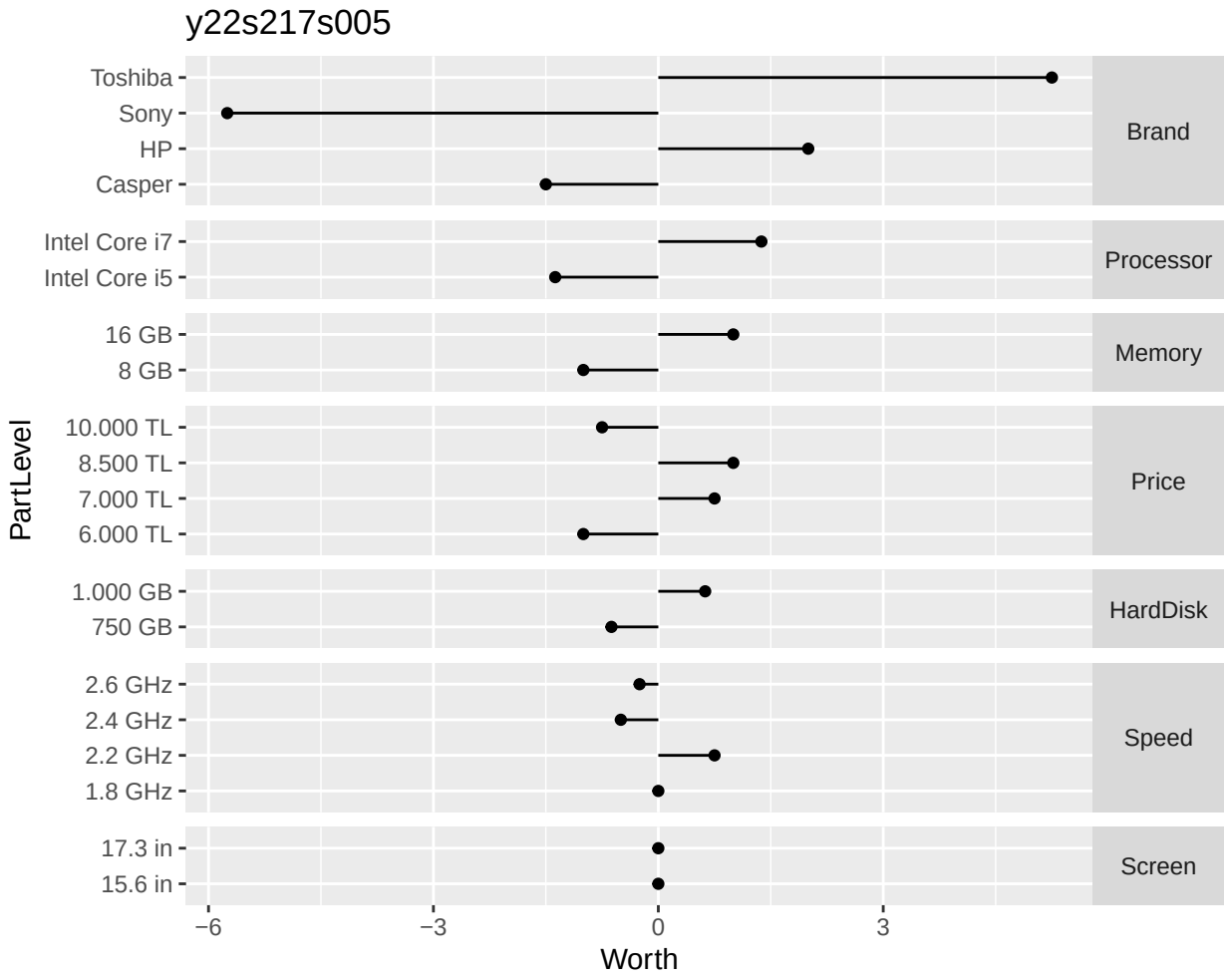

y22s216s002
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
Toshiba
Sony
HP
Casper
Brand
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
Intel Core i7
Intel Core i5
Processor
17.3 in
15.6 in
Screen
1.000 GB
750 GB
HardDisk
16 GB
8 GB
Memory
PartLevel
−6
−3
0
3
6
Worth

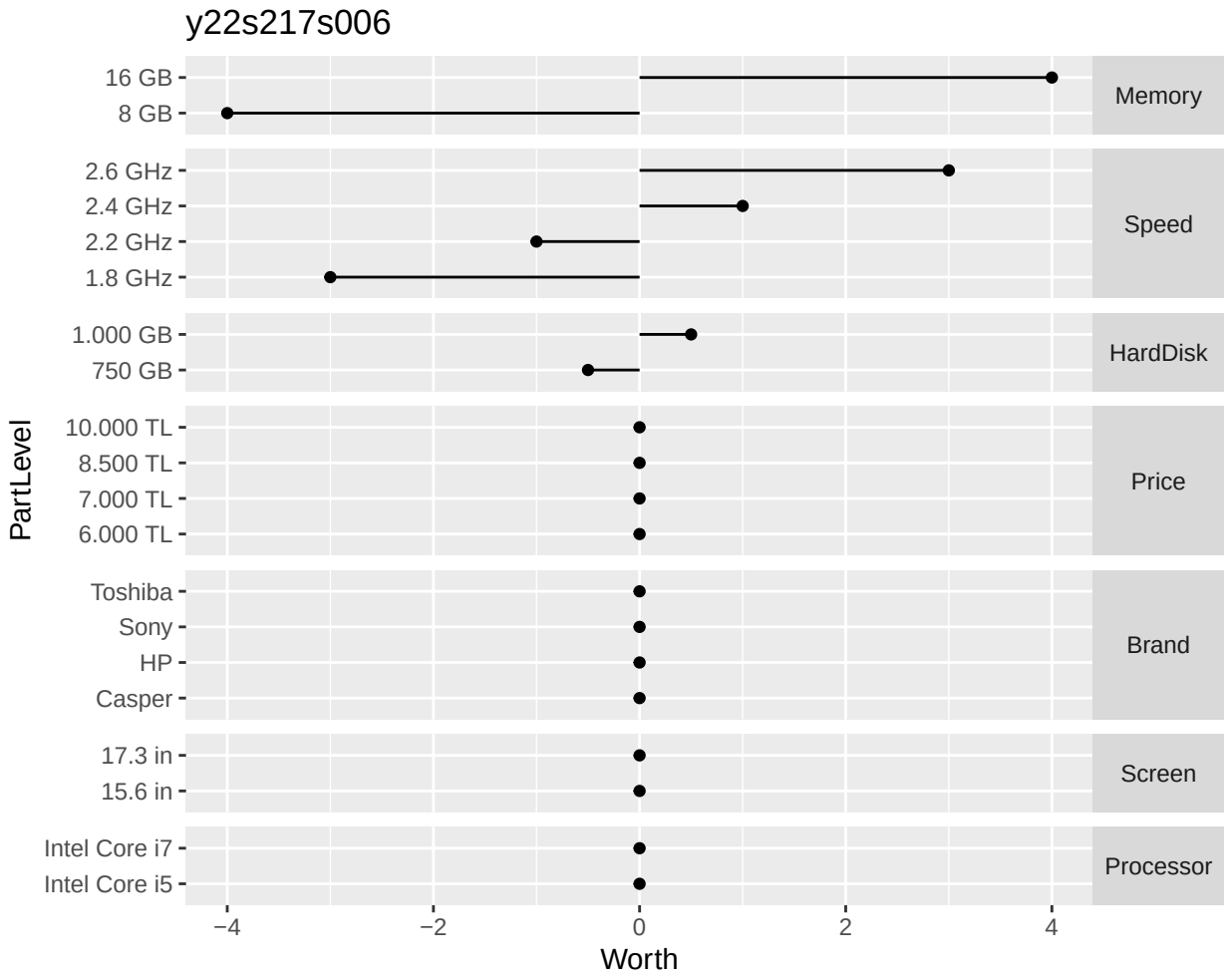

y22s216s003
16 GB
8 GB
Memory
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
Toshiba
Sony
HP
Casper
Brand
Intel Core i7
Intel Core i5
Processor
1.000 GB
750 GB
HardDisk
17.3 in
15.6 in
Screen
PartLevel
−2
0
2
Worth

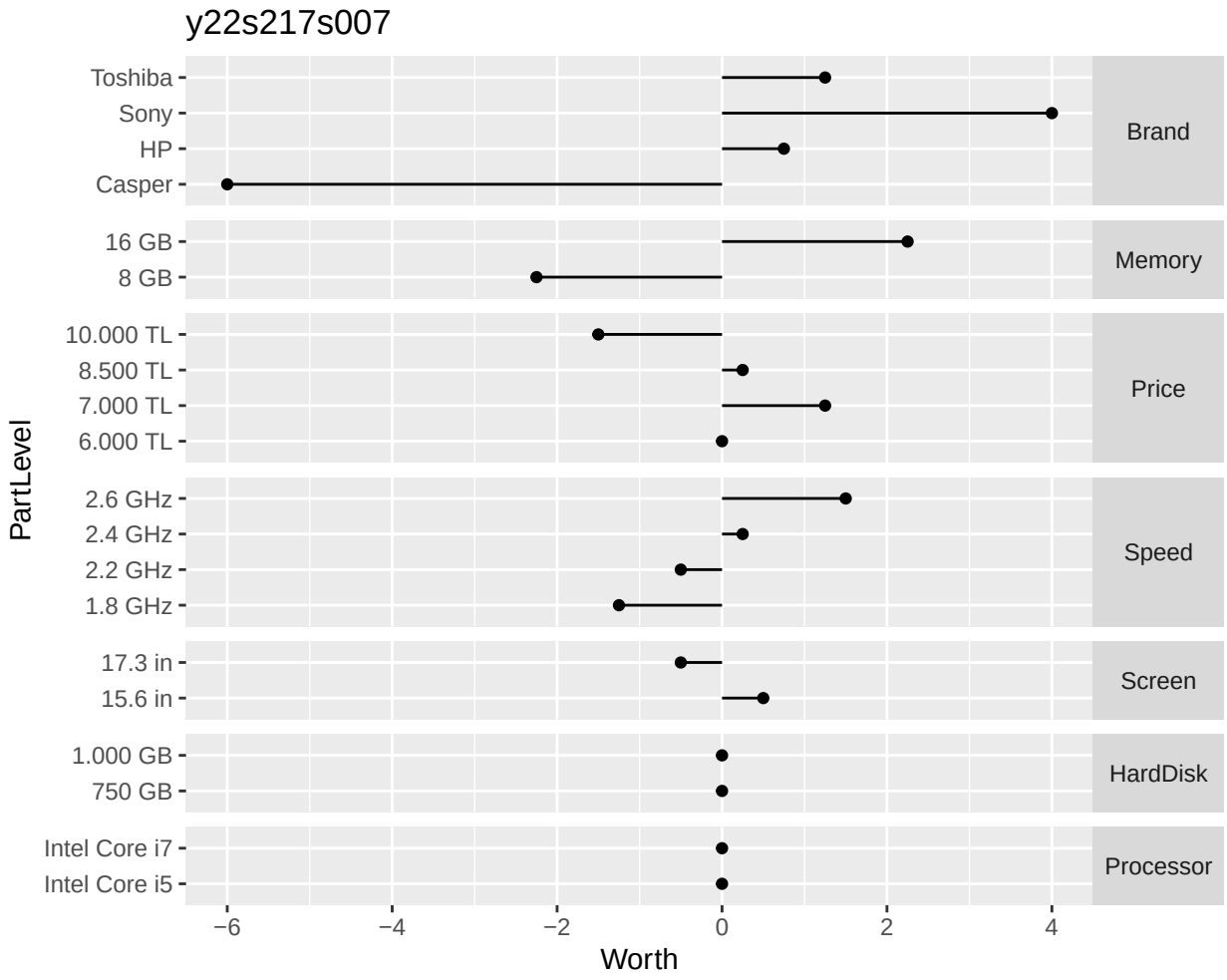

y22s217s002
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
Intel Core i7
Intel Core i5
Processor
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
Toshiba
Sony
HP
Casper
Brand
16 GB
8 GB
Memory
17.3 in
15.6 in
Screen
1.000 GB
750 GB
HardDisk
PartLevel
−6
−4
−2
0
2
Worth

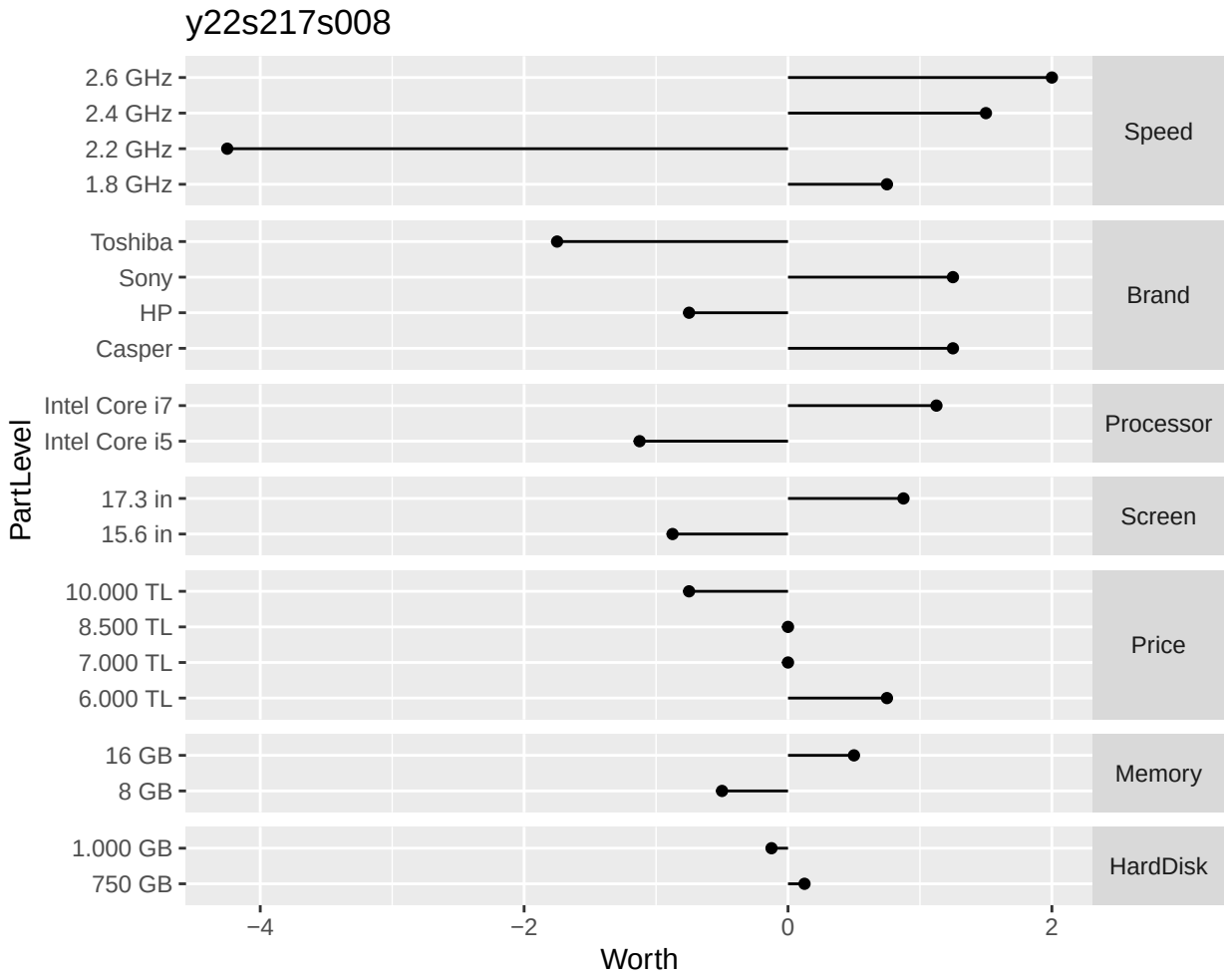

y22s217s003
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
Toshiba
Sony
HP
Casper
Brand
1.000 GB
750 GB
HardDisk
Intel Core i7
Intel Core i5
Processor
16 GB
8 GB
Memory
17.3 in
15.6 in
Screen
PartLevel
−4
−2
0
2
4
Worth

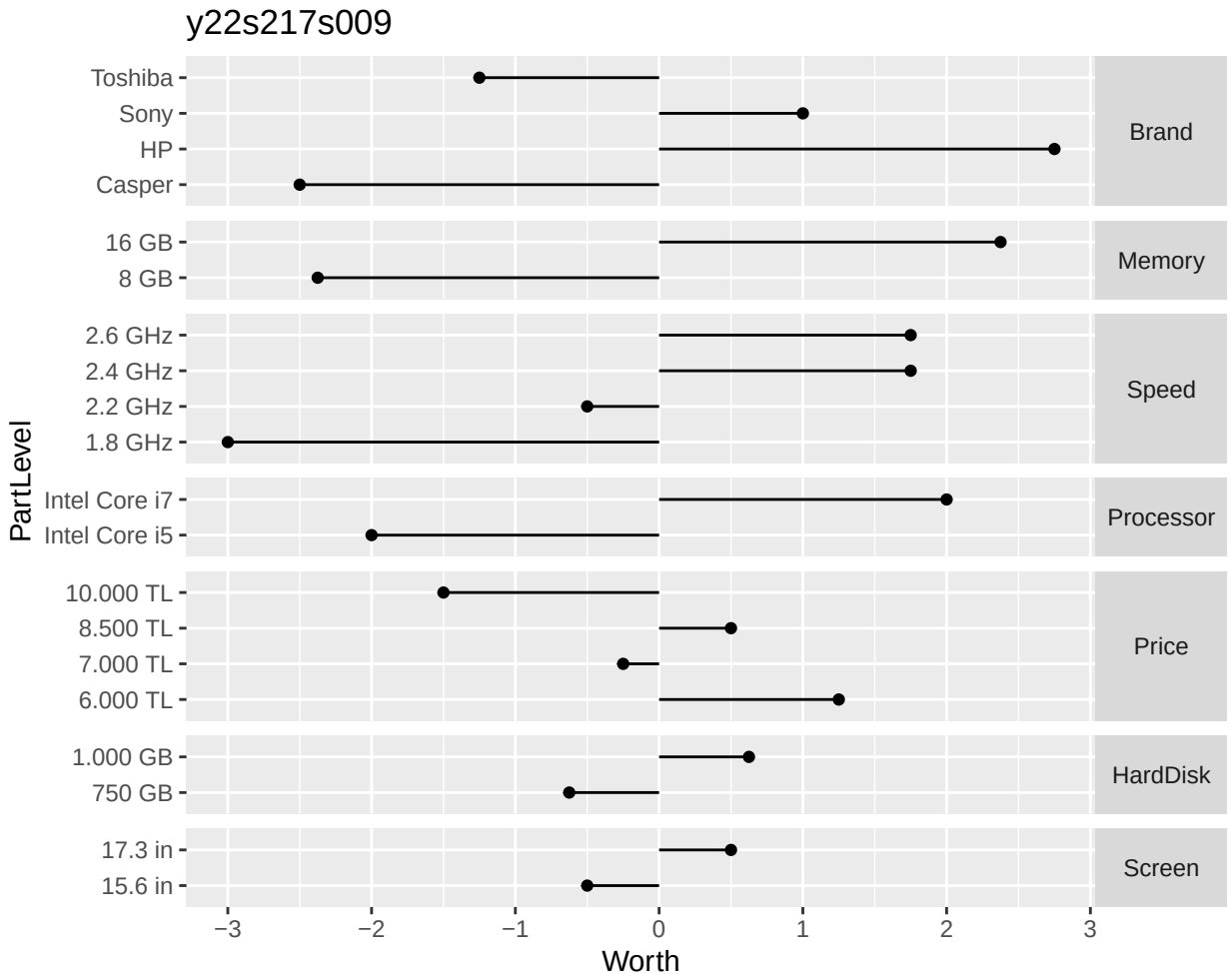

y22s217s004
Intel Core i7
Intel Core i5
Processor
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
17.3 in
15.6 in
Screen
Toshiba
Sony
HP
Casper
Brand
1.000 GB
750 GB
HardDisk
16 GB
8 GB
Memory
PartLevel
−4
−2
0
2
4
Worth

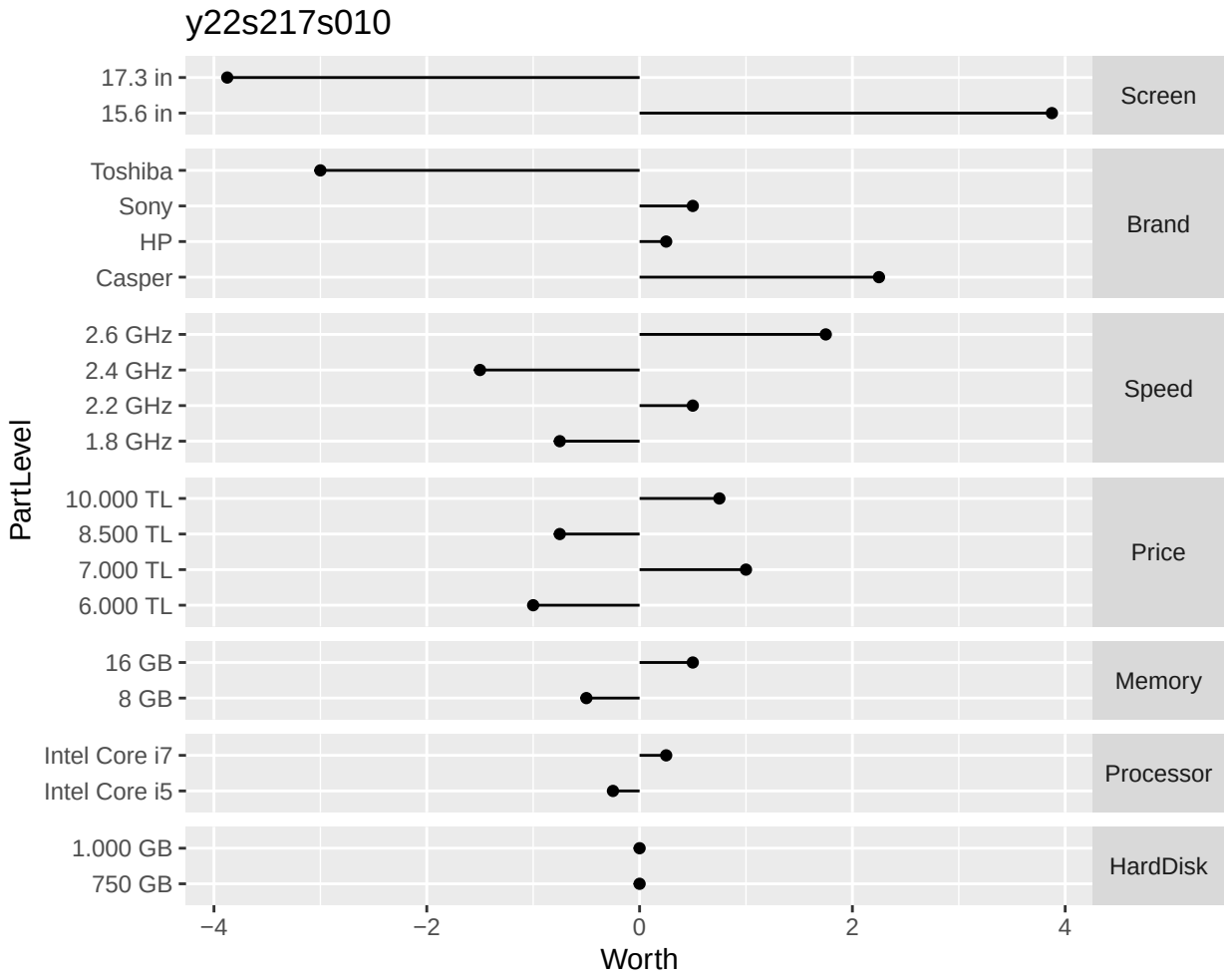

y22s217s005
Toshiba
Sony
HP
Casper
Brand
Intel Core i7
Intel Core i5
Processor
16 GB
8 GB
Memory
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
1.000 GB
750 GB
HardDisk
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
17.3 in
15.6 in
Screen
PartLevel
−6
−3
0
3
Worth

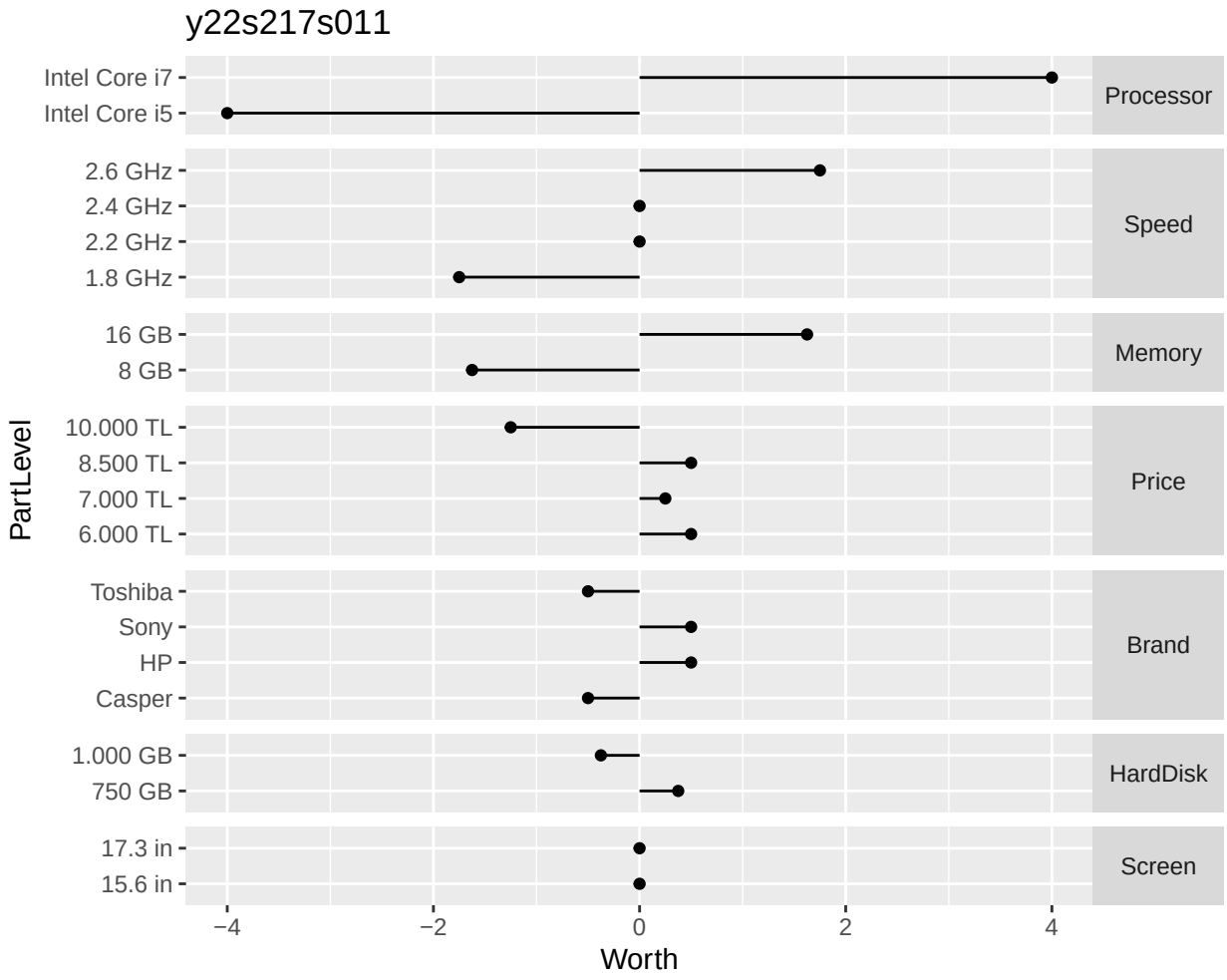

y22s217s006
16 GB
8 GB
Memory
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
1.000 GB
750 GB
HardDisk
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
Toshiba
Sony
HP
Casper
Brand
17.3 in
15.6 in
Screen
Intel Core i7
Intel Core i5
Processor
PartLevel
−4
−2
0
2
4
Worth

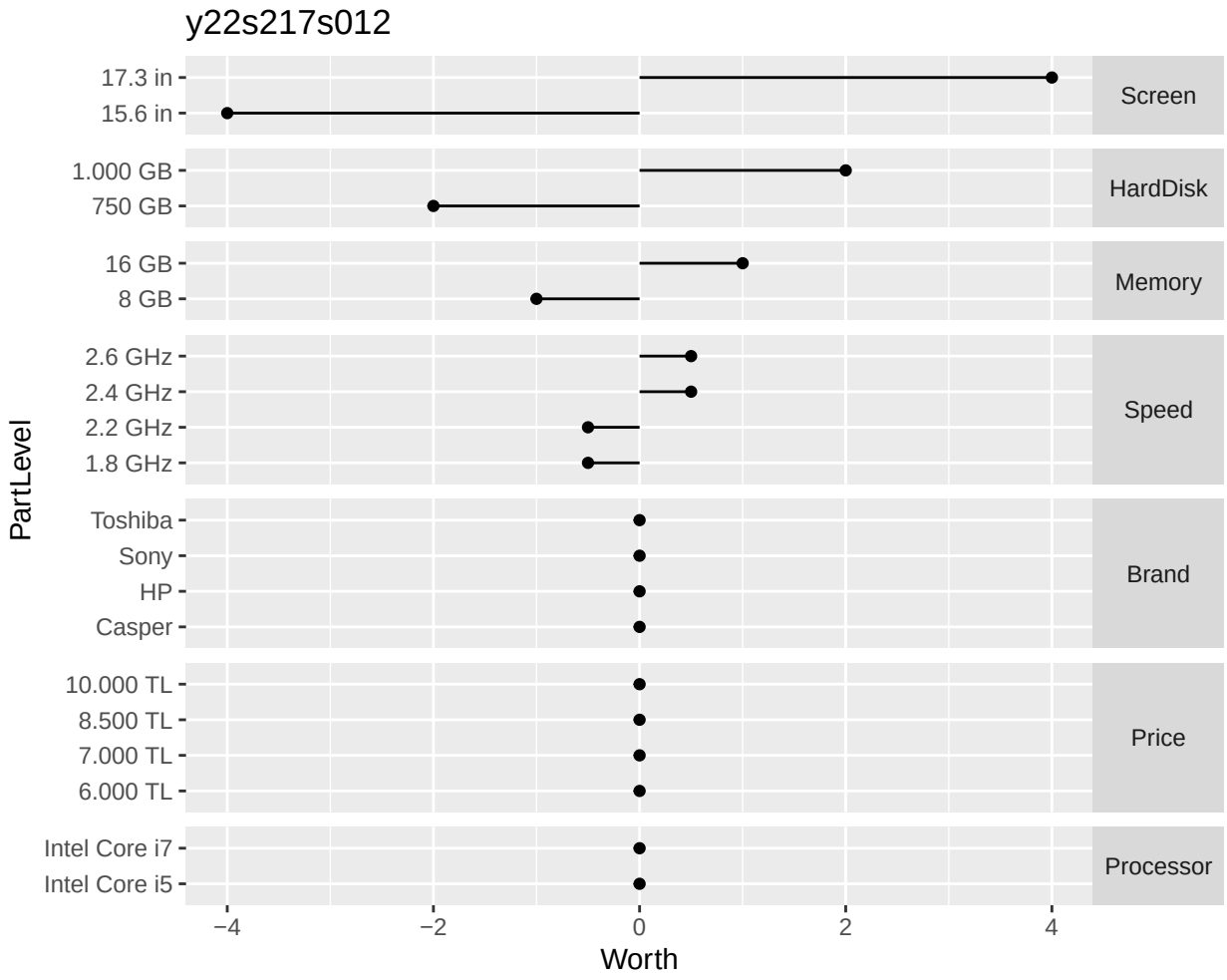

y22s217s007
Toshiba
Sony
HP
Casper
Brand
16 GB
8 GB
Memory
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
17.3 in
15.6 in
Screen
1.000 GB
750 GB
HardDisk
Intel Core i7
Intel Core i5
Processor
PartLevel
−6
−4
−2
0
2
4
Worth

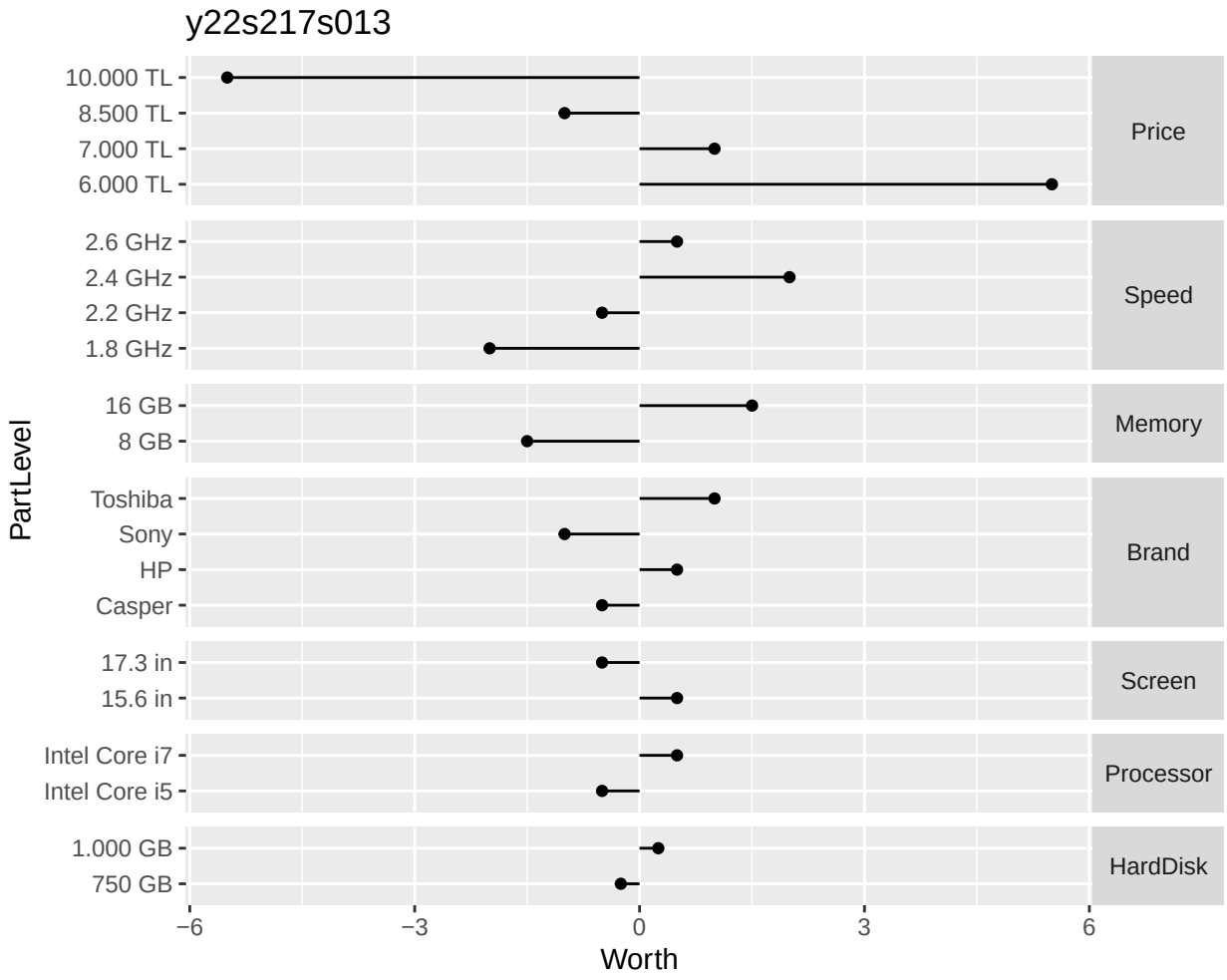

y22s217s008
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
Toshiba
Sony
HP
Casper
Brand
Intel Core i7
Intel Core i5
Processor
17.3 in
15.6 in
Screen
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
16 GB
8 GB
Memory
1.000 GB
750 GB
HardDisk
PartLevel
−4
−2
0
2
Worth

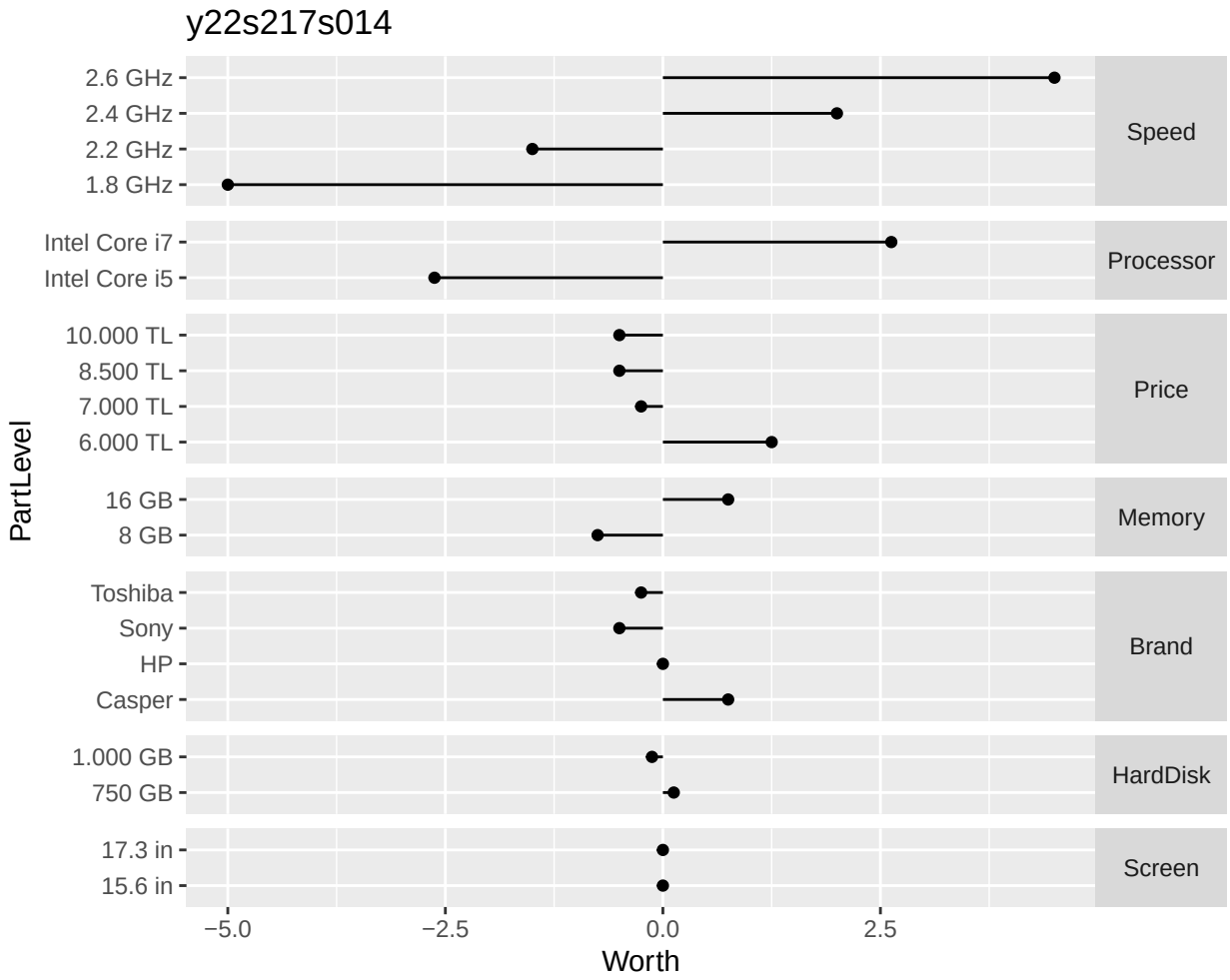

y22s217s009
Toshiba
Sony
HP
Casper
Brand
16 GB
8 GB
Memory
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
Intel Core i7
Intel Core i5
Processor
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
1.000 GB
750 GB
HardDisk
17.3 in
15.6 in
Screen
PartLevel
−3
−2
−1
0
1
2
3
Worth

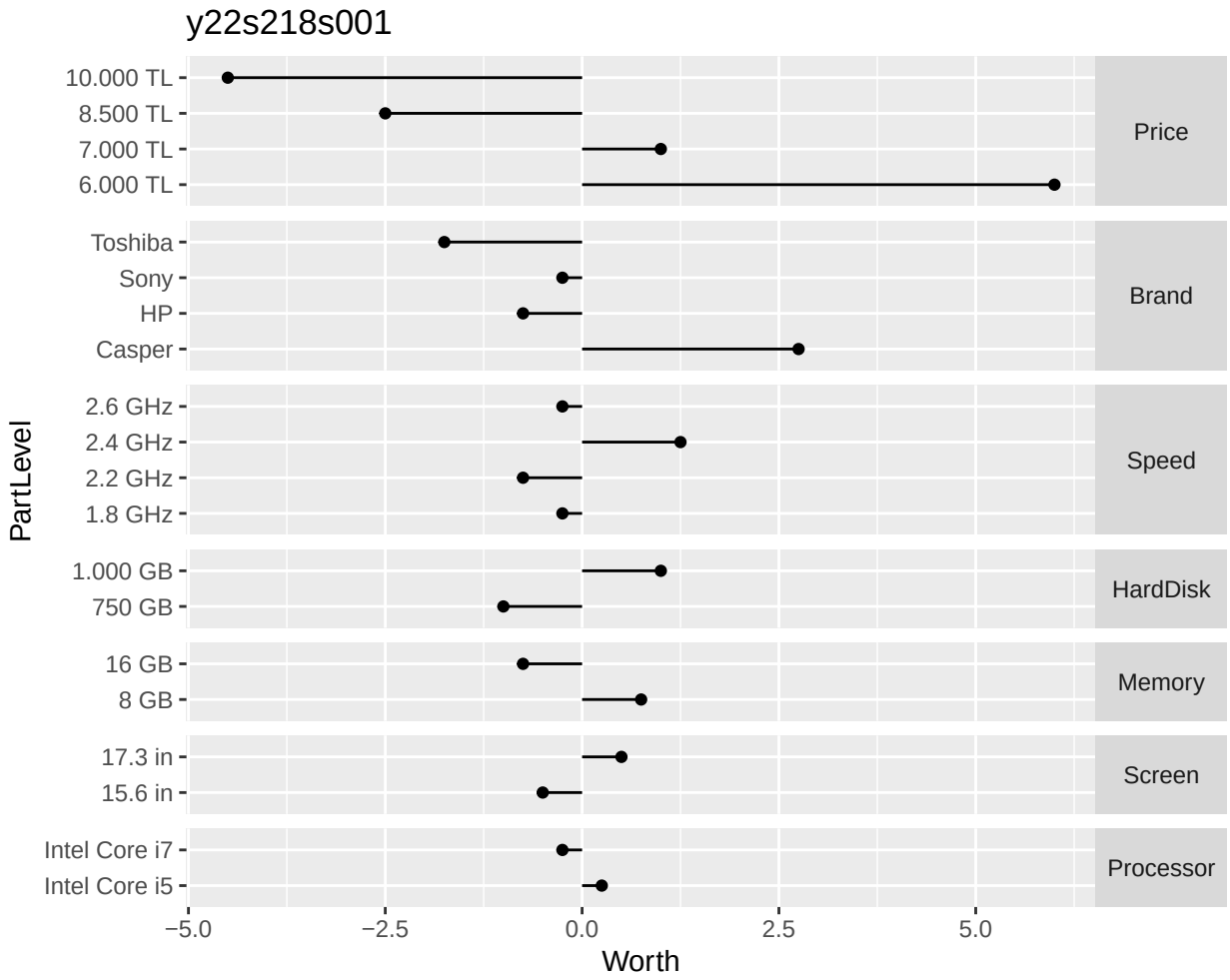

y22s217s010
17.3 in
15.6 in
Screen
Toshiba
Sony
HP
Casper
Brand
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
16 GB
8 GB
Memory
Intel Core i7
Intel Core i5
Processor
1.000 GB
750 GB
HardDisk
PartLevel
−4
−2
0
2
4
Worth

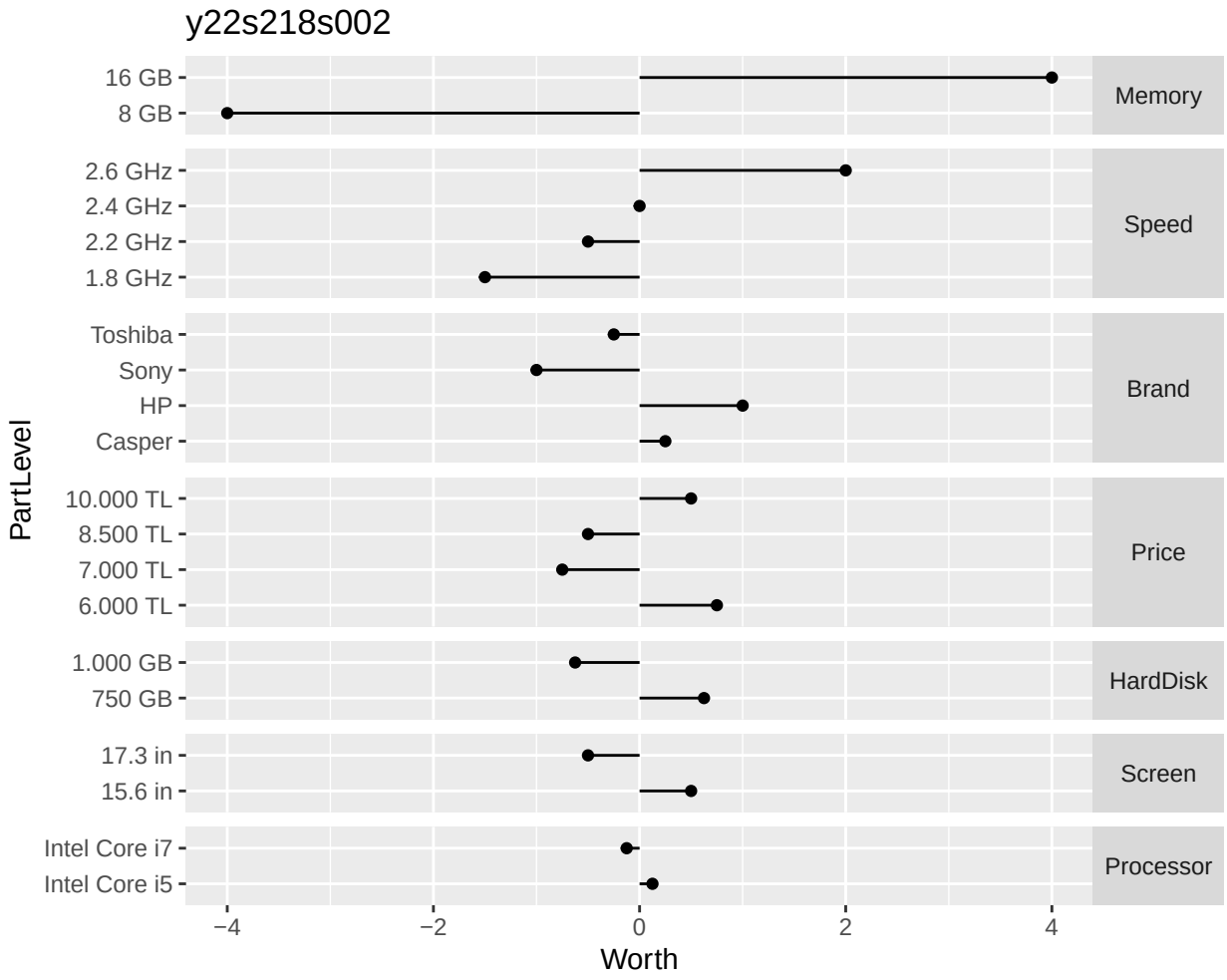

y22s217s011
Intel Core i7
Intel Core i5
Processor
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
16 GB
8 GB
Memory
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
Toshiba
Sony
HP
Casper
Brand
1.000 GB
750 GB
HardDisk
17.3 in
15.6 in
Screen
PartLevel
−4
−2
0
2
4
Worth

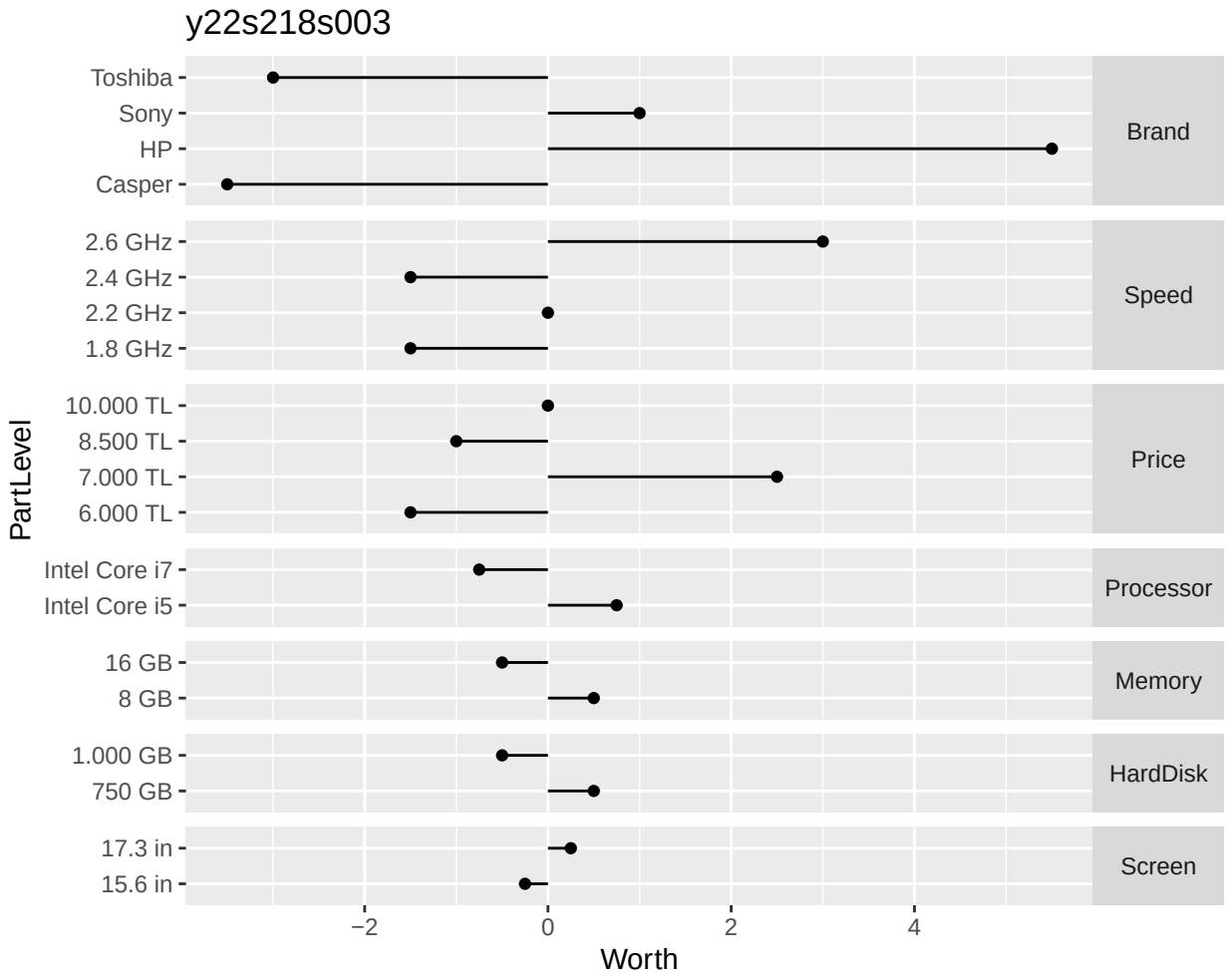

y22s217s012
17.3 in
15.6 in
Screen
1.000 GB
750 GB
HardDisk
16 GB
8 GB
Memory
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
Toshiba
Sony
HP
Casper
Brand
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
Intel Core i7
Intel Core i5
Processor
PartLevel
−4
−2
0
2
4
Worth

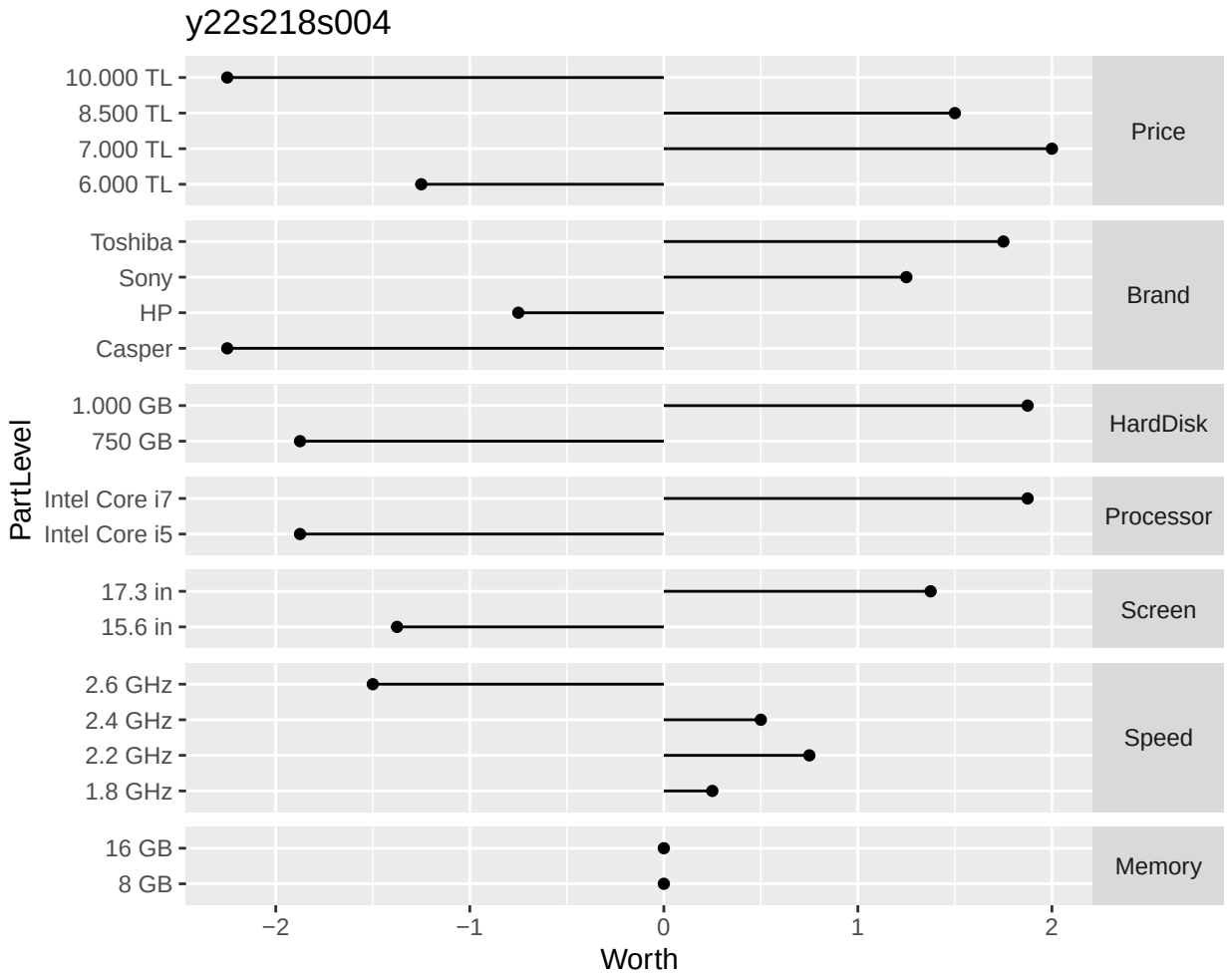

y22s217s013
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
16 GB
8 GB
Memory
Toshiba
Sony
HP
Casper
Brand
17.3 in
15.6 in
Screen
Intel Core i7
Intel Core i5
Processor
1.000 GB
750 GB
HardDisk
PartLevel
−6
−3
0
3
6
Worth

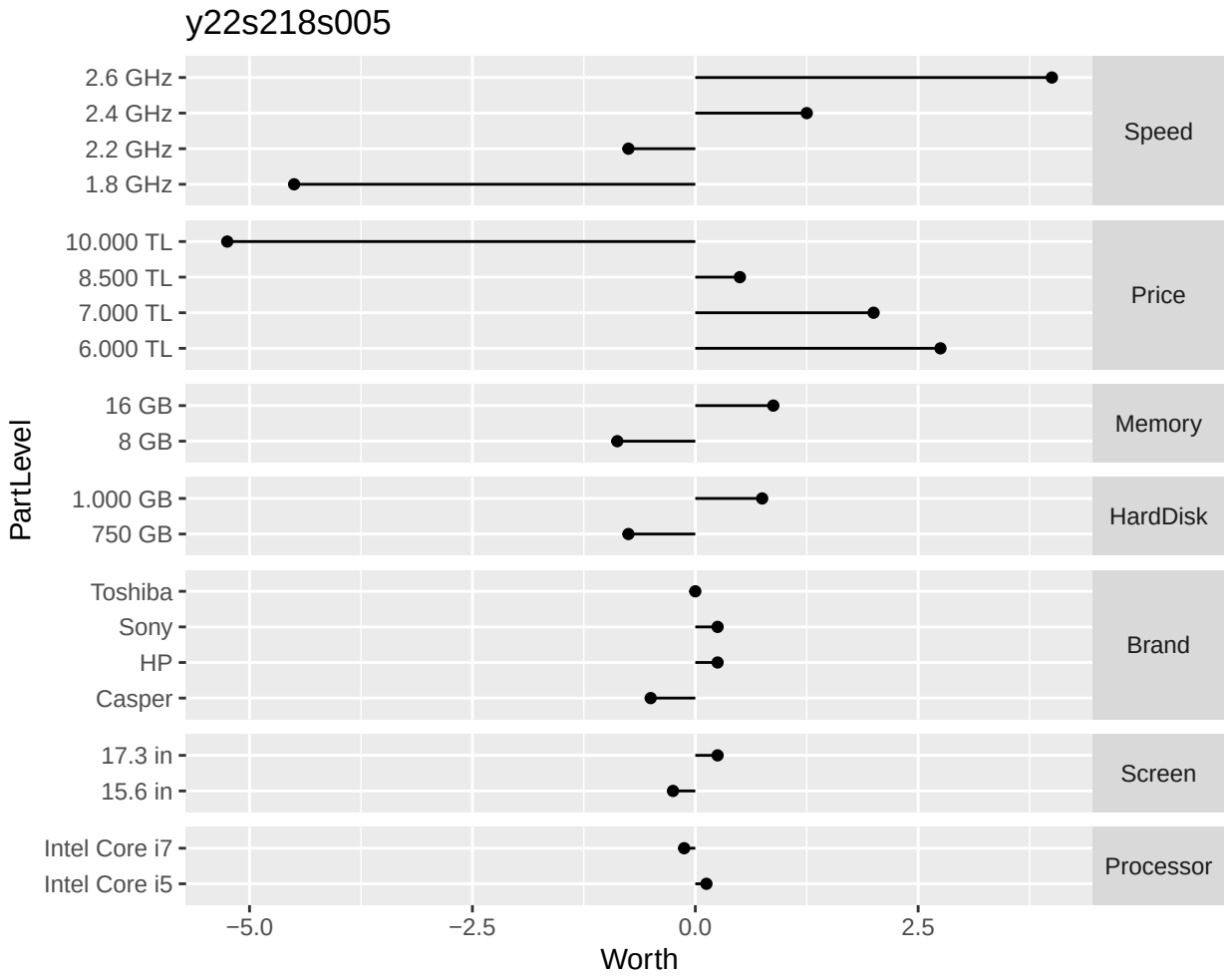

y22s217s014
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
Intel Core i7
Intel Core i5
Processor
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
16 GB
8 GB
Memory
Toshiba
Sony
HP
Casper
Brand
1.000 GB
750 GB
HardDisk
17.3 in
15.6 in
Screen
PartLevel
−5.0
−2.5
0.0
2.5
Worth

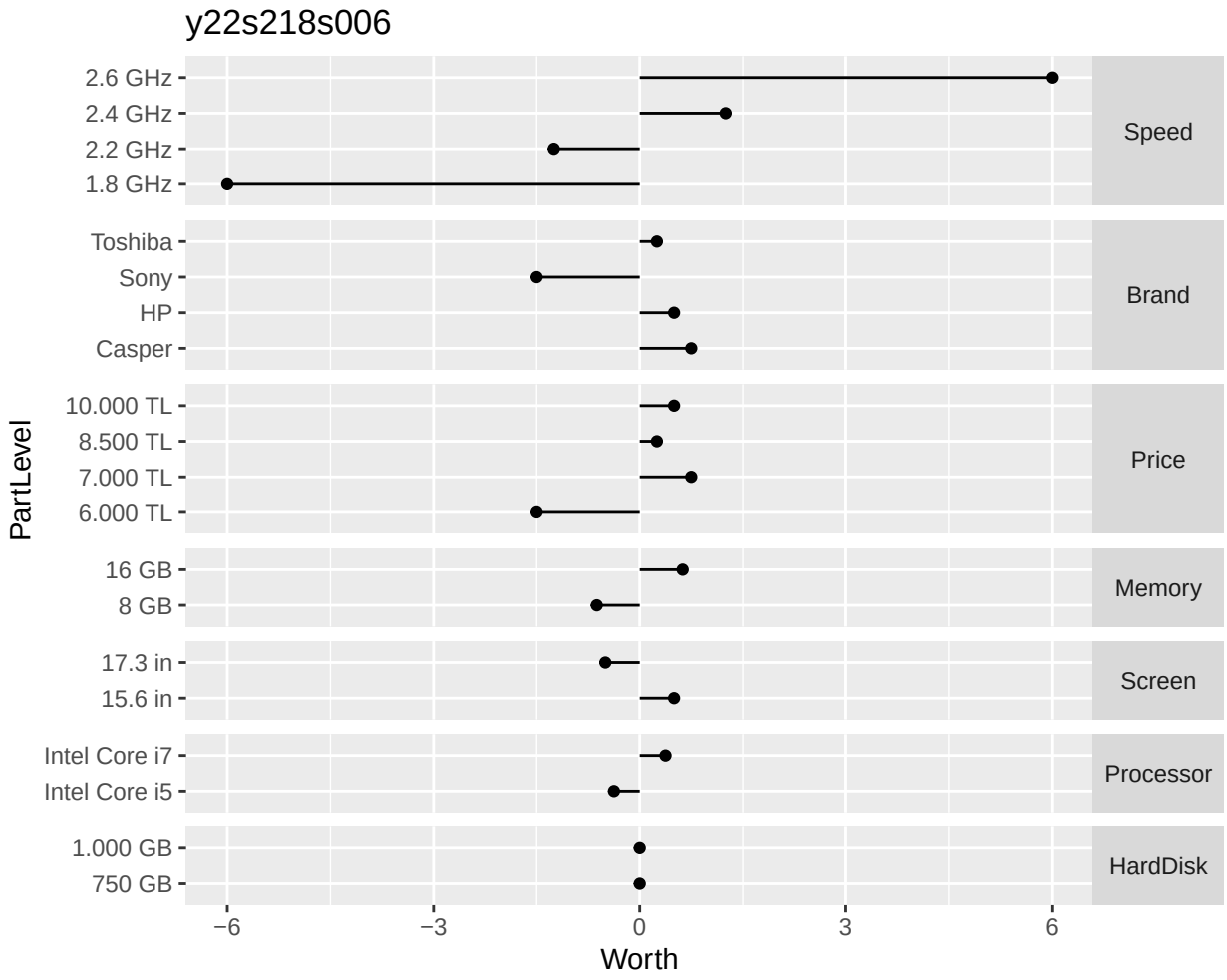

y22s218s001
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
Toshiba
Sony
HP
Casper
Brand
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
1.000 GB
750 GB
HardDisk
16 GB
8 GB
Memory
17.3 in
15.6 in
Screen
Intel Core i7
Intel Core i5
Processor
PartLevel
−5.0
−2.5
0.0
2.5
5.0
Worth

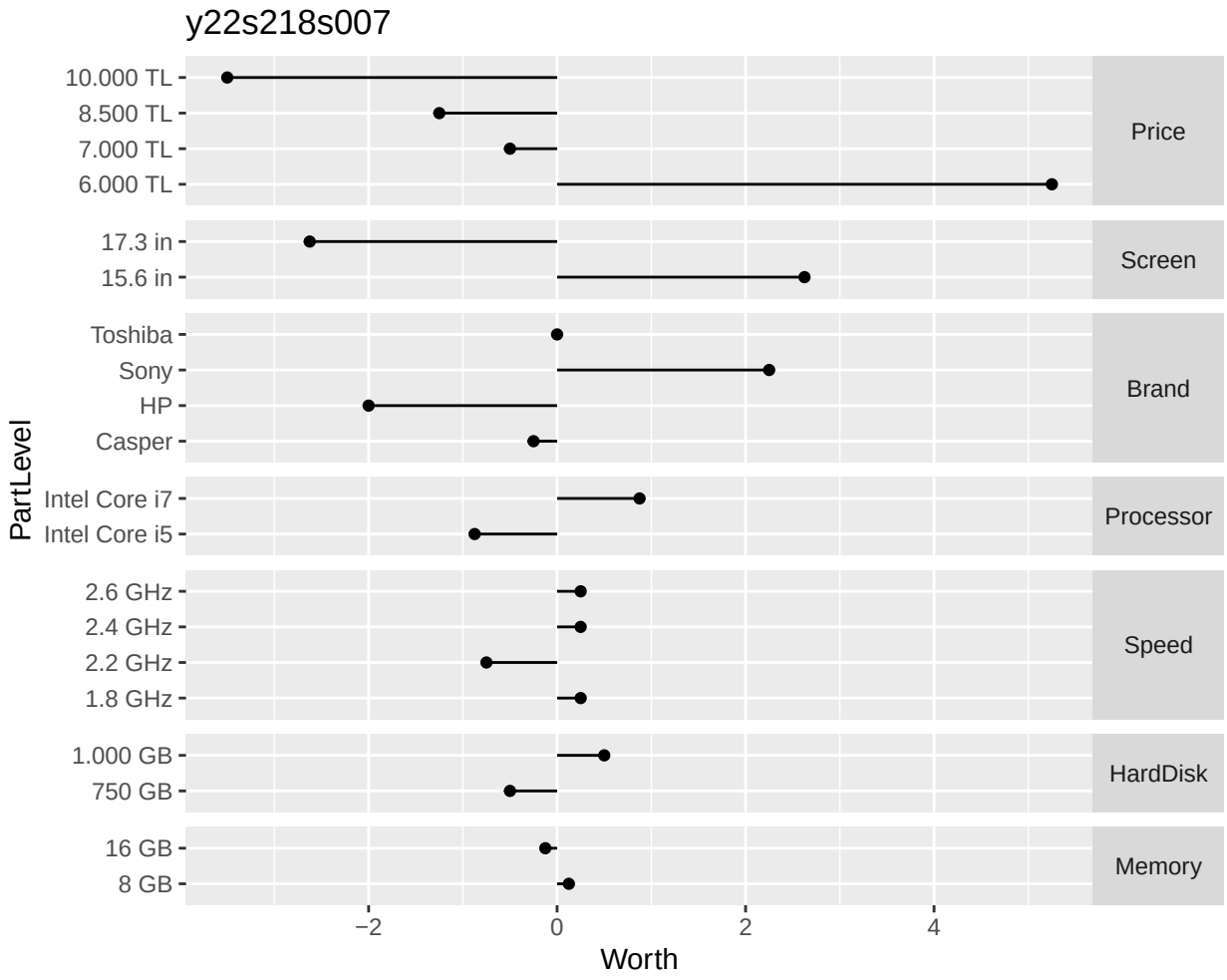

y22s218s002
16 GB
8 GB
Memory
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
Toshiba
Sony
HP
Casper
Brand
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
1.000 GB
750 GB
HardDisk
17.3 in
15.6 in
Screen
Intel Core i7
Intel Core i5
Processor
PartLevel
−4
−2
0
2
4
Worth

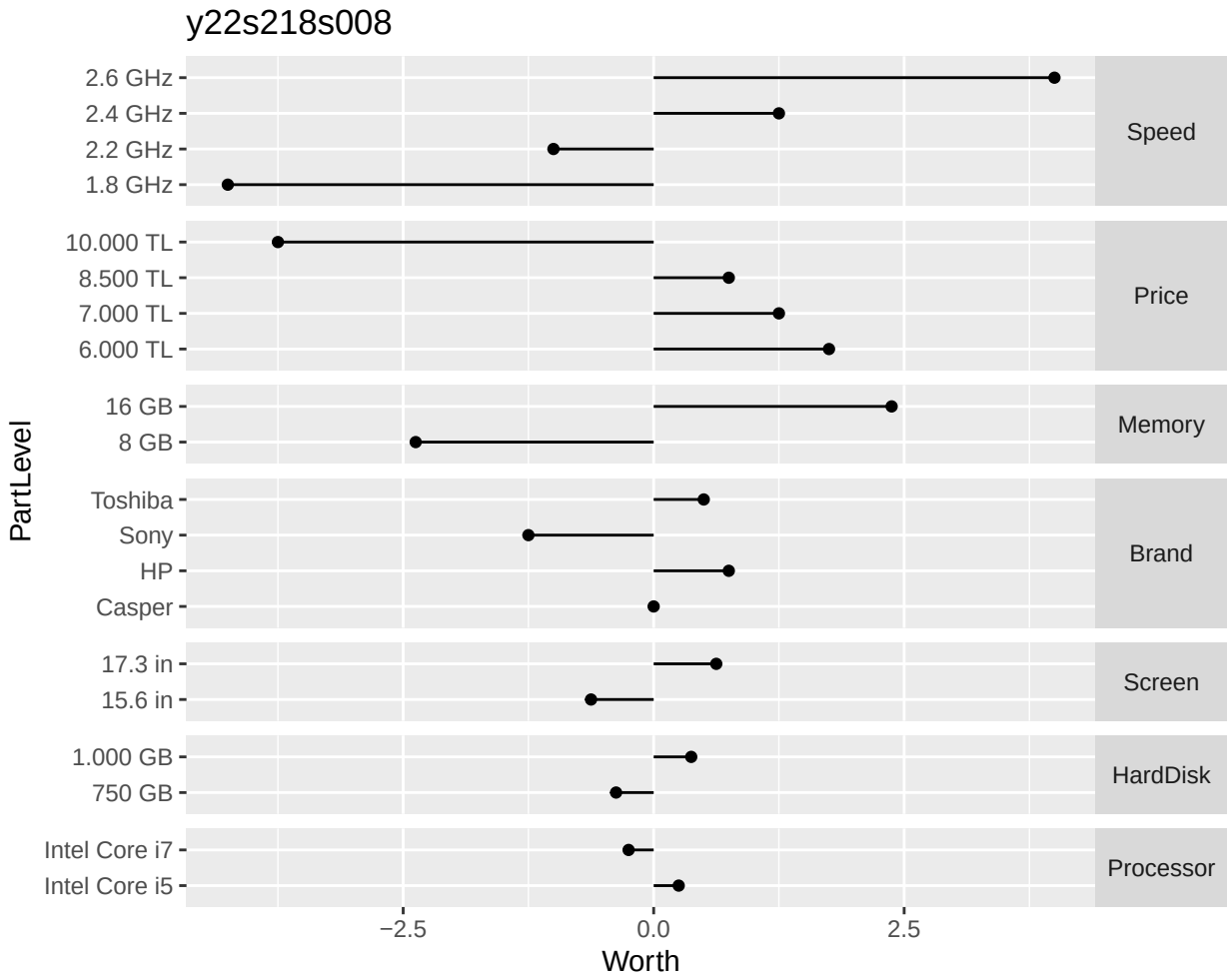

y22s218s003
Toshiba
Sony
HP
Casper
Brand
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
Intel Core i7
Intel Core i5
Processor
16 GB
8 GB
Memory
1.000 GB
750 GB
HardDisk
17.3 in
15.6 in
Screen
PartLevel
−2
0
2
4
Worth

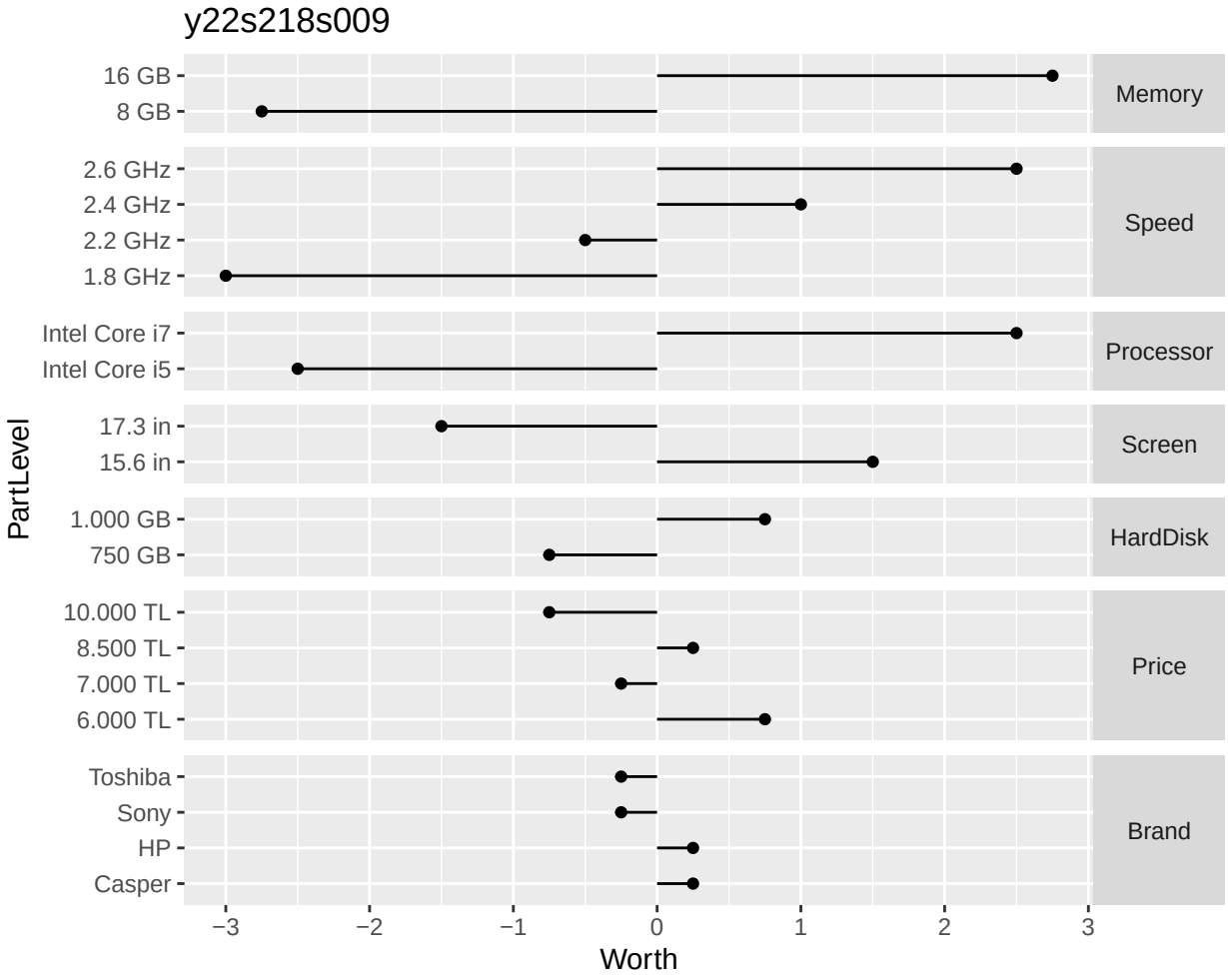

y22s218s004
PartLevel
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
Toshiba
Sony
HP
Casper
Brand
1.000 GB
750 GB
HardDisk
Intel Core i7
Intel Core i5
Processor
17.3 in
15.6 in
Screen
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
16 GB
8 GB
Memory
−2
−1
0
1
2
Worth

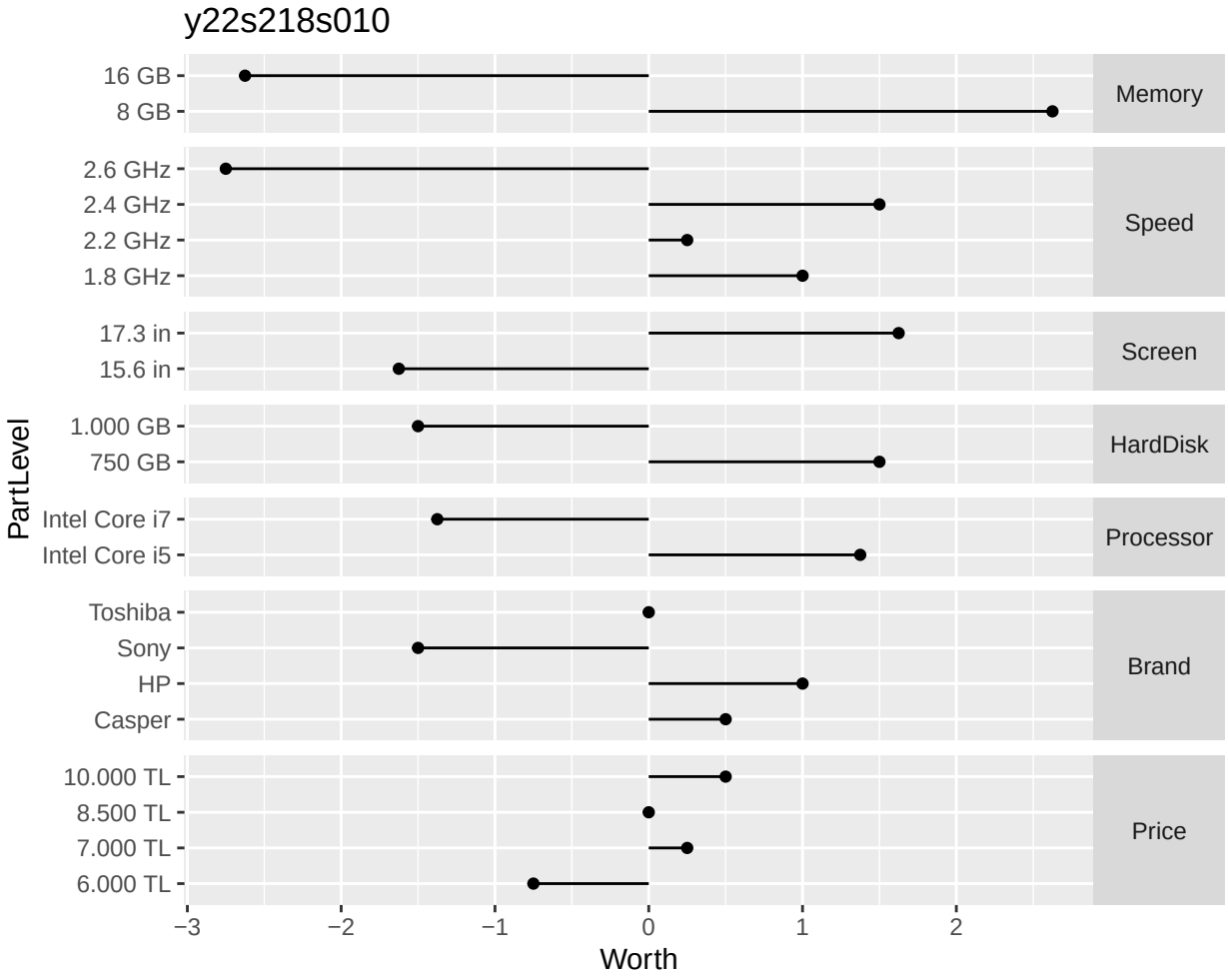

y22s218s005
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
16 GB
8 GB
Memory
1.000 GB
750 GB
HardDisk
Toshiba
Sony
HP
Casper
Brand
17.3 in
15.6 in
Screen
Intel Core i7
Intel Core i5
Processor
PartLevel
−5.0
−2.5
0.0
2.5
Worth

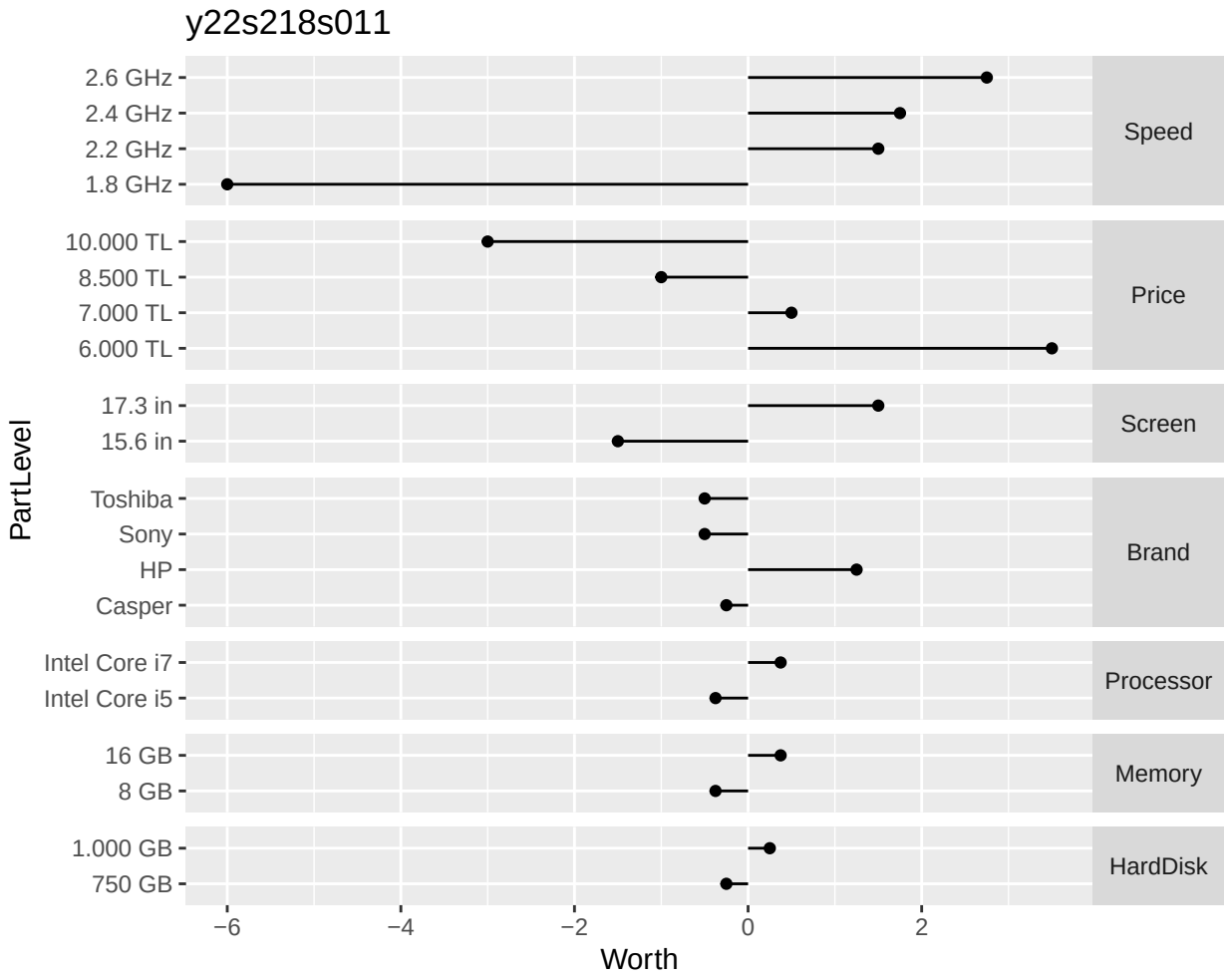

y22s218s006
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
Toshiba
Sony
HP
Casper
Brand
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
16 GB
8 GB
Memory
17.3 in
15.6 in
Screen
Intel Core i7
Intel Core i5
Processor
1.000 GB
750 GB
HardDisk
PartLevel
−6
−3
0
3
6
Worth

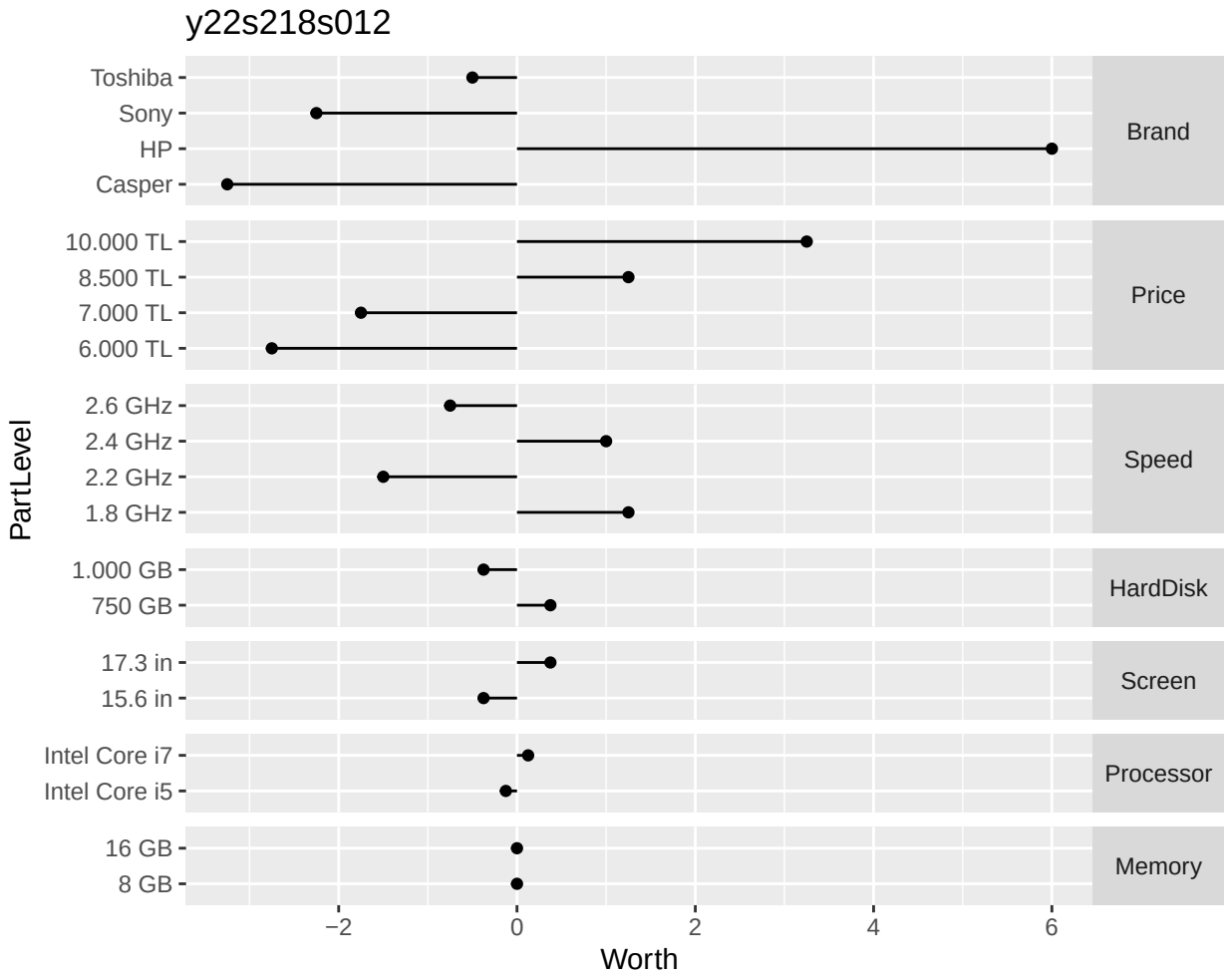

y22s218s007
PartLevel
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
17.3 in
15.6 in
Screen
Toshiba
Sony
HP
Casper
Brand
Intel Core i7
Intel Core i5
Processor
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
1.000 GB
750 GB
HardDisk
16 GB
8 GB
Memory
−2
0
2
4
Worth

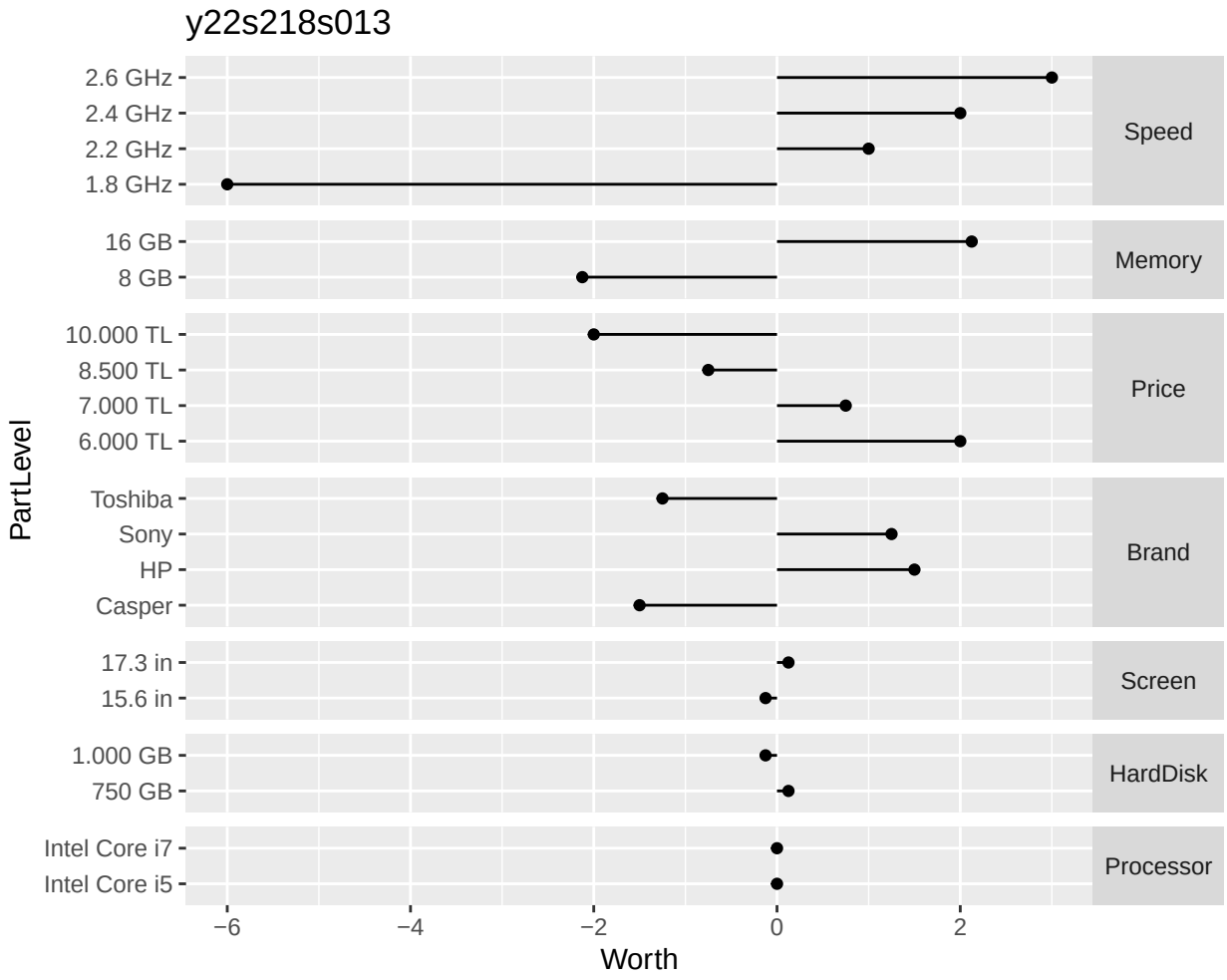

y22s218s008
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
16 GB
8 GB
Memory
Toshiba
Sony
HP
Casper
Brand
17.3 in
15.6 in
Screen
1.000 GB
750 GB
HardDisk
Intel Core i7
Intel Core i5
Processor
PartLevel
−2.5
0.0
2.5
Worth

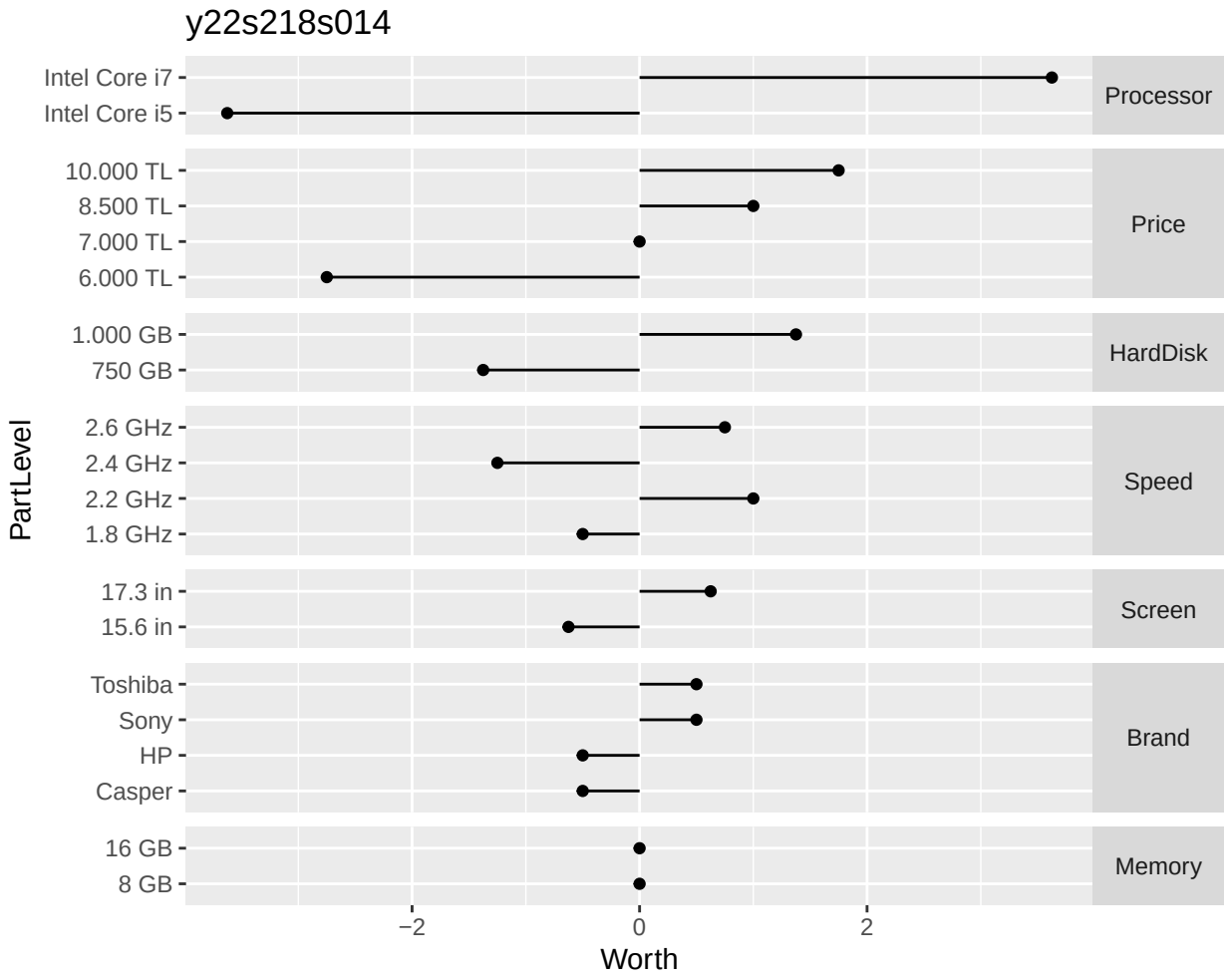

y22s218s009
16 GB
8 GB
Memory
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
Intel Core i7
Intel Core i5
Processor
17.3 in
15.6 in
Screen
1.000 GB
750 GB
HardDisk
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
Toshiba
Sony
HP
Casper
Brand
PartLevel
−3
−2
−1
0
1
2
3
Worth

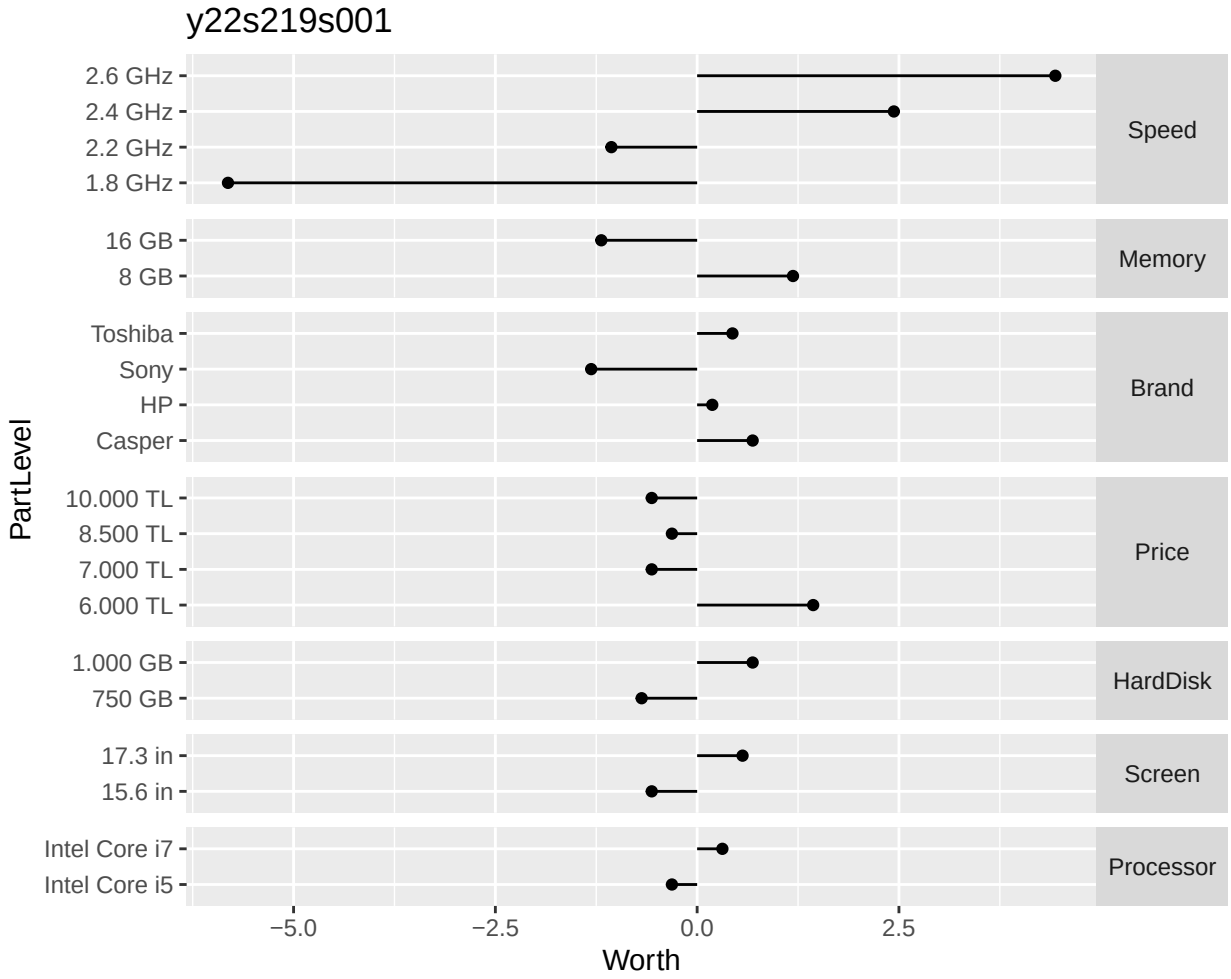

y22s218s010
16 GB
8 GB
Memory
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
17.3 in
15.6 in
Screen
1.000 GB
750 GB
HardDisk
Intel Core i7
Intel Core i5
Processor
Toshiba
Sony
HP
Casper
Brand
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
PartLevel
−3
−2
−1
0
1
2
Worth

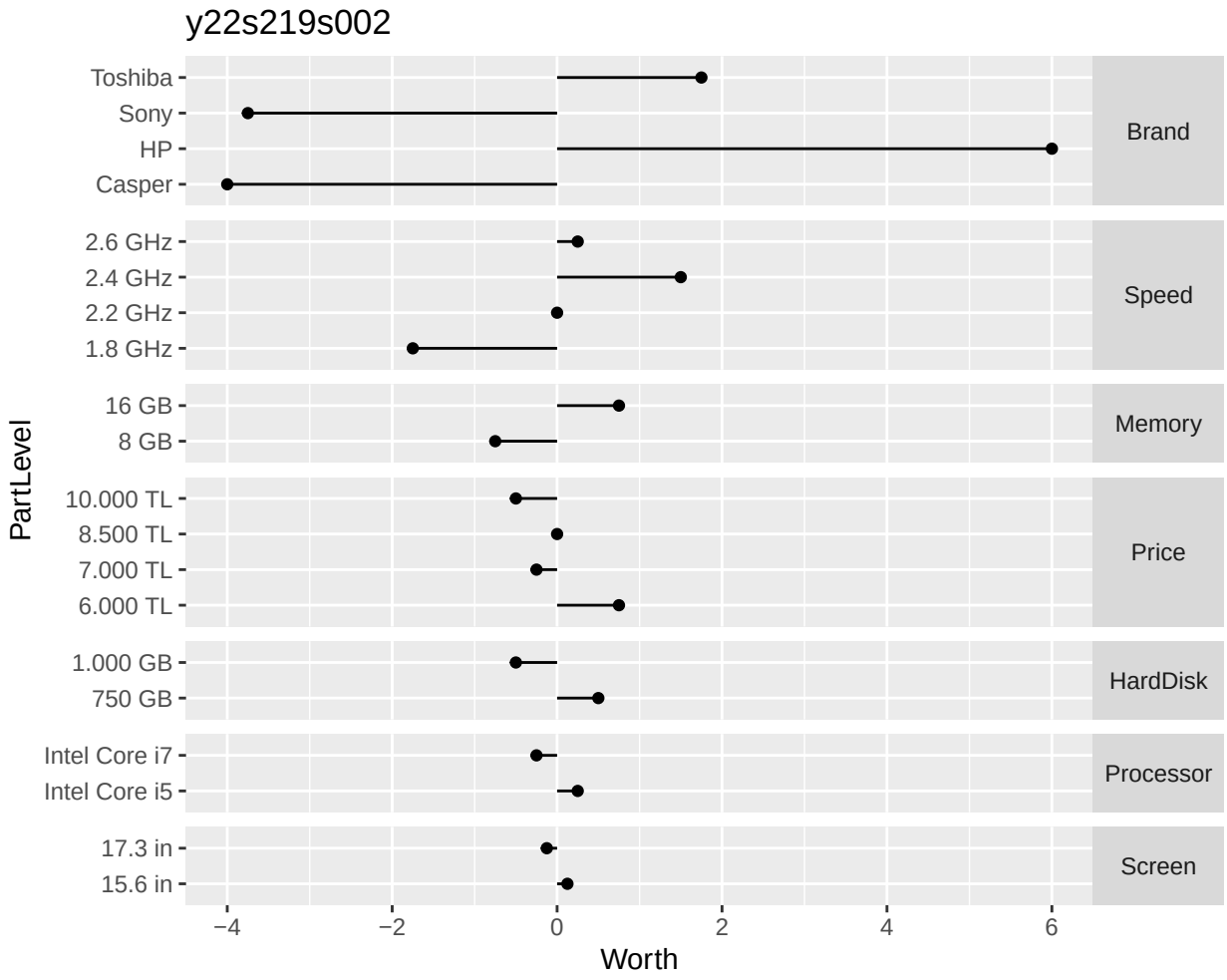

y22s218s011
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
17.3 in
15.6 in
Screen
Toshiba
Sony
HP
Casper
Brand
Intel Core i7
Intel Core i5
Processor
16 GB
8 GB
Memory
1.000 GB
750 GB
HardDisk
PartLevel
−6
−4
−2
0
2
Worth

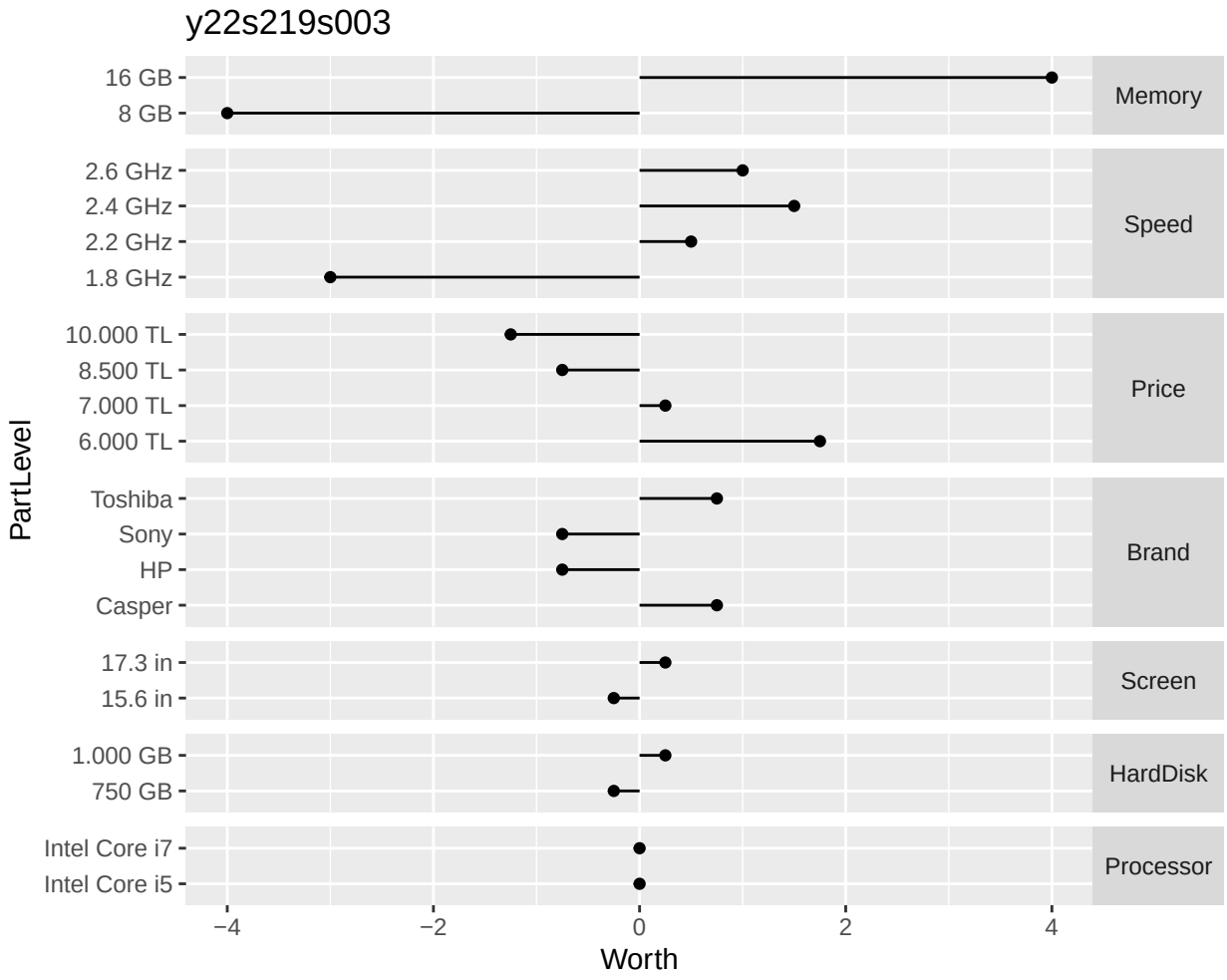

y22s218s012
Toshiba
Sony
HP
Casper
Brand
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
1.000 GB
750 GB
HardDisk
17.3 in
15.6 in
Screen
Intel Core i7
Intel Core i5
Processor
16 GB
8 GB
Memory
PartLevel
−2
0
2
4
6
Worth

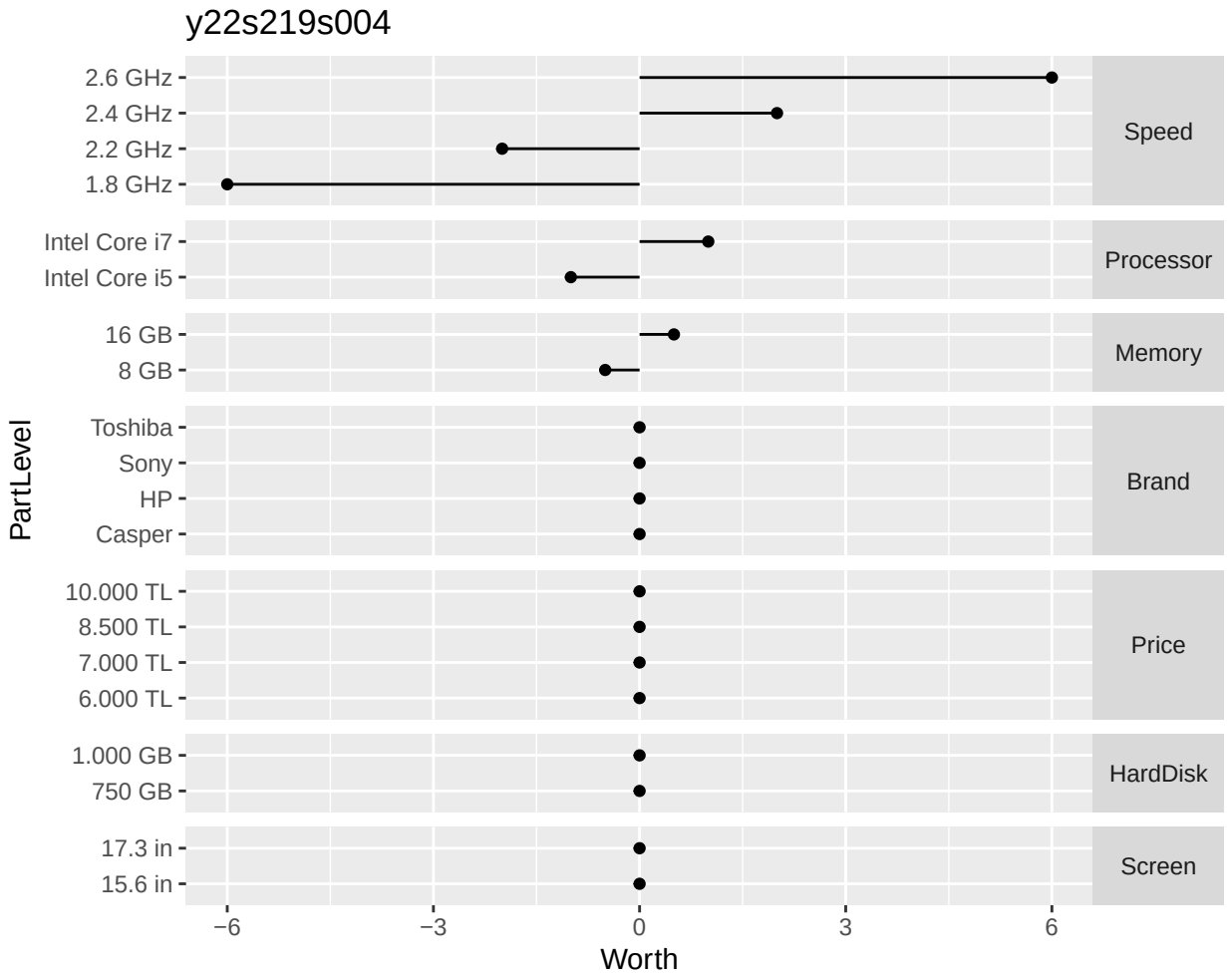

y22s218s013
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
16 GB
8 GB
Memory
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
Toshiba
Sony
HP
Casper
Brand
17.3 in
15.6 in
Screen
1.000 GB
750 GB
HardDisk
Intel Core i7
Intel Core i5
Processor
PartLevel
−6
−4
−2
0
2
Worth

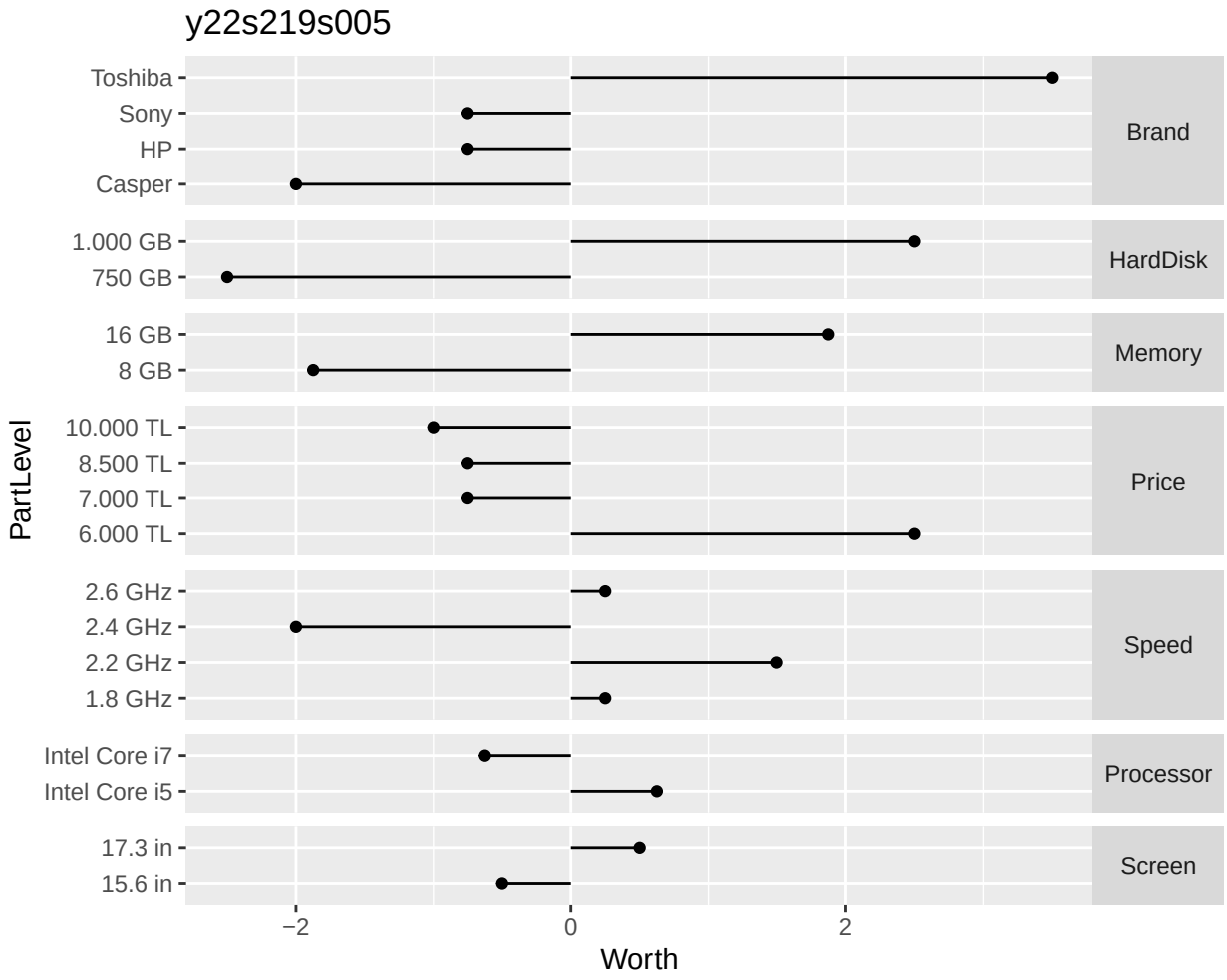

y22s218s014
Intel Core i7
Intel Core i5
Processor
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
1.000 GB
750 GB
HardDisk
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
17.3 in
15.6 in
Screen
Toshiba
Sony
HP
Casper
Brand
16 GB
8 GB
Memory
PartLevel
−2
0
2
Worth


y22s219s001
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
16 GB
8 GB
Memory
Toshiba
Sony
HP
Casper
Brand
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
1.000 GB
750 GB
HardDisk
17.3 in
15.6 in
Screen
Intel Core i7
Intel Core i5
Processor
PartLevel
−5.0
−2.5
0.0
2.5
Worth


y22s219s002
Toshiba
Sony
HP
Casper
Brand
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
16 GB
8 GB
Memory
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
1.000 GB
750 GB
HardDisk
Intel Core i7
Intel Core i5
Processor
17.3 in
15.6 in
Screen
PartLevel
−4
−2
0
2
4
6
Worth


y22s219s003
16 GB
8 GB
Memory
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
Toshiba
Sony
HP
Casper
Brand
17.3 in
15.6 in
Screen
1.000 GB
750 GB
HardDisk
Intel Core i7
Intel Core i5
Processor
PartLevel
−4
−2
0
2
4
Worth


y22s219s004
2.6 GHz
2.4 GHz
2.2 GHz
1.8 GHz
Speed
Intel Core i7
Intel Core i5
Processor
16 GB
8 GB
Memory
Toshiba
Sony
HP
Casper
Brand
10.000 TL
8.500 TL
7.000 TL
6.000 TL
Price
1.000 GB
750 GB
HardDisk
17.3 in
15.6 in
Screen
PartLevel
−6
−3
0
3
6
Worth

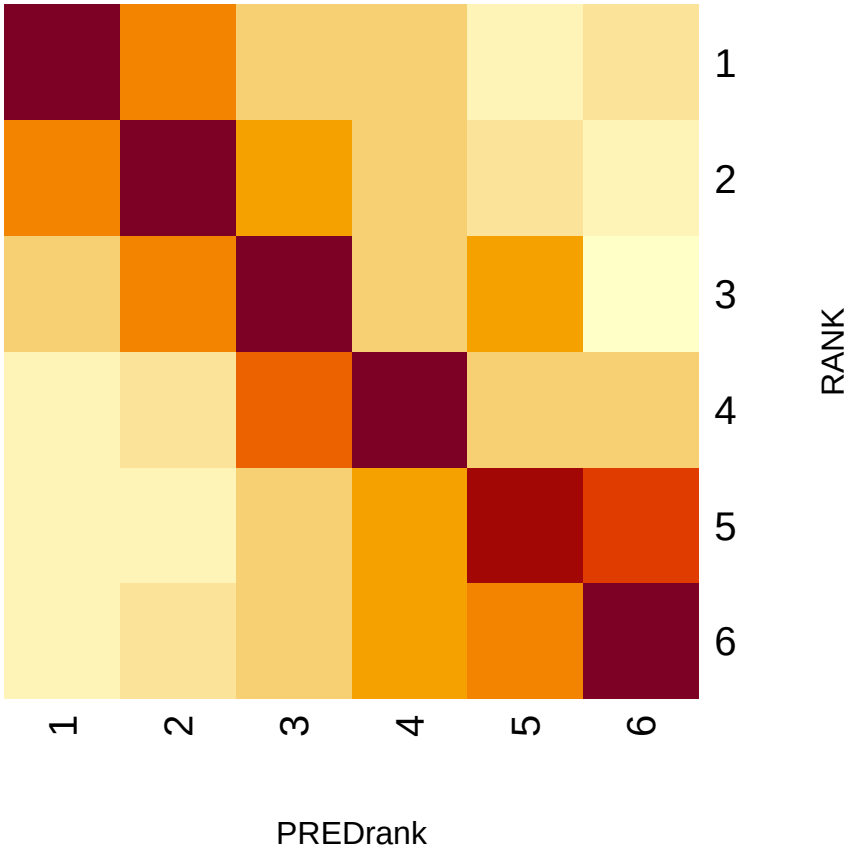


Validation of market model

```
confusion <- market_df %>%
  select(id, test_pred) %>%
  unnest(test_pred) %>%
  xtabs(~ RANK + PREDrank, .)

confusion
```

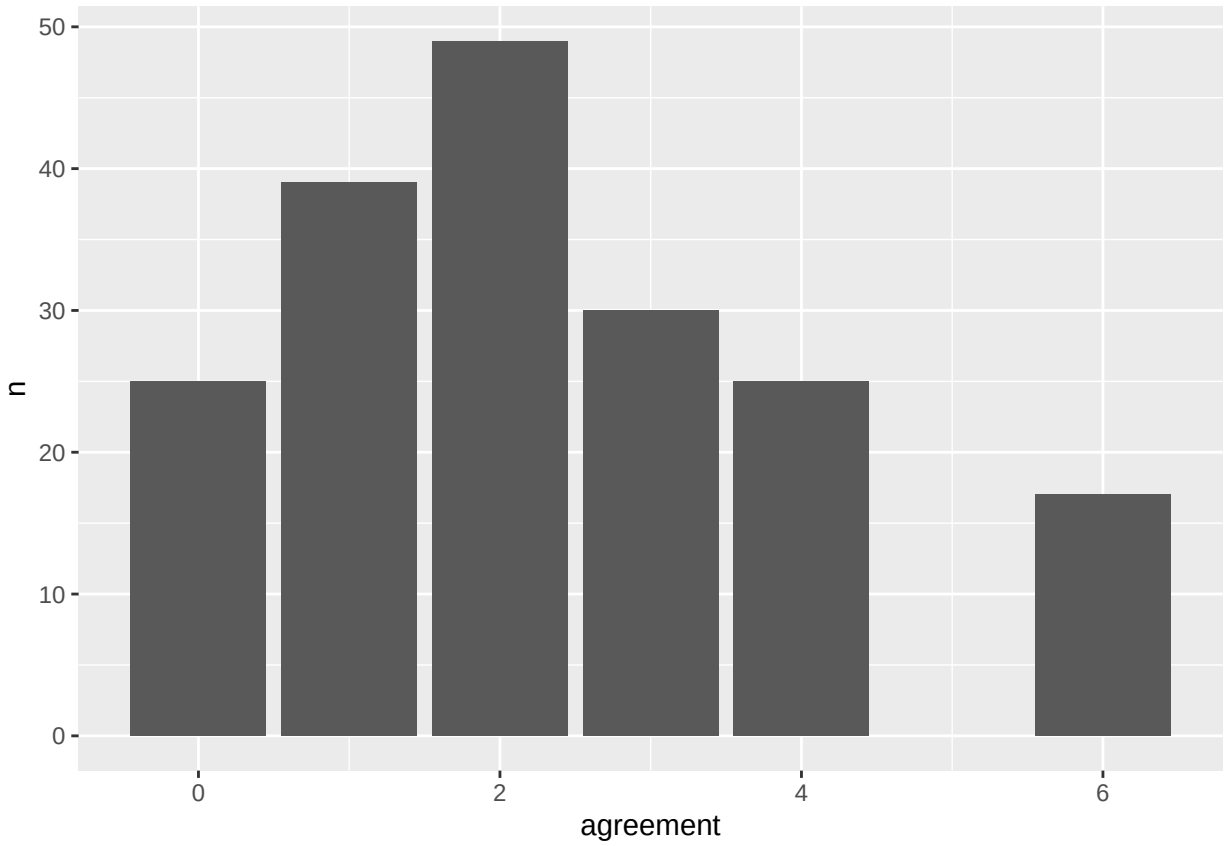
```
##     PREDrank
## RANK   1   2   3   4   5   6
##    1 100  43  16  16   2   8
##    2  38  68  32  22  16   9
##    3  24  38  57  24  32  10
##    4  13  18  43  59  26  26
##    5   4   6  18  32  69  56
##    6   6  12  19  32  40  76
```

```
confnames <- dimnames(confusion) %>% names()
heatmap(confusion, Rowv = NA, Colv = NA, revC = TRUE, xlab = confnames[2], ylab = confnames[1])
```



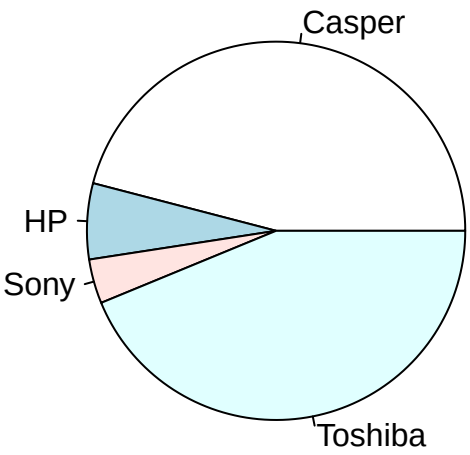

```
# confusion %>%
#   as_tibble() %>%
#   rename(Observed = RANK, Predicted = PRED) %>%
#   ggplot(aes(Predicted, Observed)) +
#   geom_tile(aes(fill=n)) +
#   scale_fill_gradient(low = "yellow", high = "red") +
#   coord_fixed()
```

```
market_df %>%
  select(id, test_pred) %>%
  unnest(test_pred) %>%
  group_by(id) %>%
  summarize(agreement = sum(RANK == PREDrank), .groups = "drop") %>%
  arrange(desc(agreement)) %>%
  count(agreement) %>%
  ggplot(aes(agreement, n)) +
  geom_col()
```



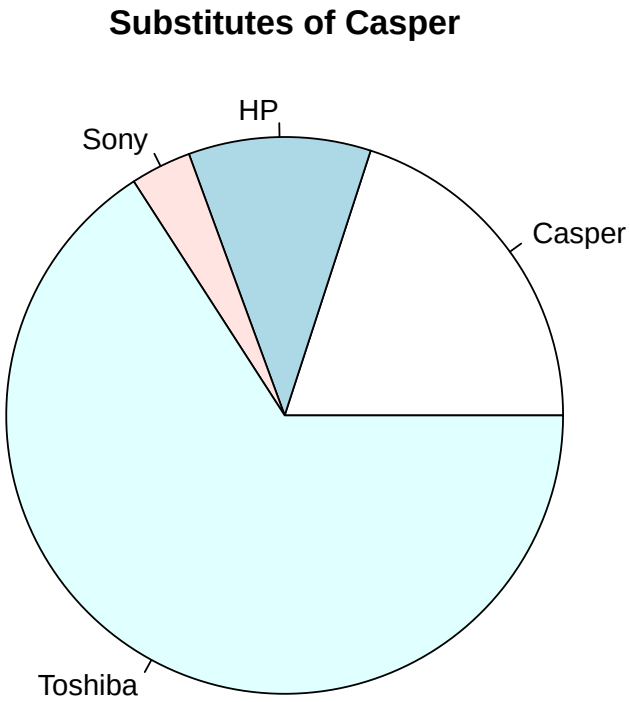
Market shares

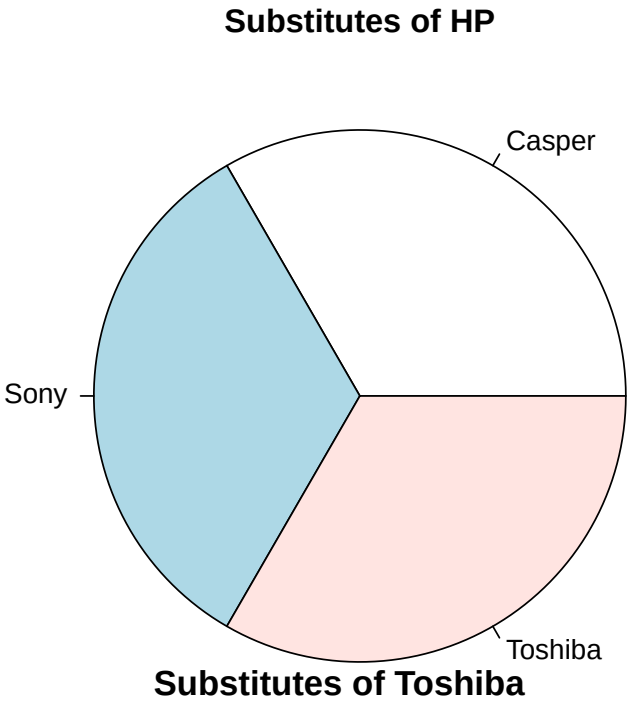
```
mshares <- market_df %>%
  select(test_data) %>%
  unnest(test_data) %>%
  filter(RANK==1) %>%
  count(Brand) %>%
  column_to_rownames(var = "Brand") %>%
  as.matrix %>%
  drop() %>%
  prop.table() %>%
  `*`(100)

mshares %>%
  pie(main="Market shares")
```

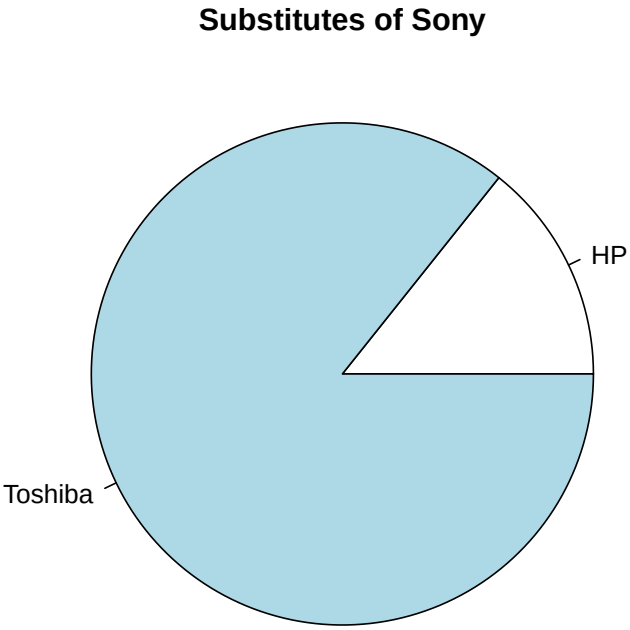


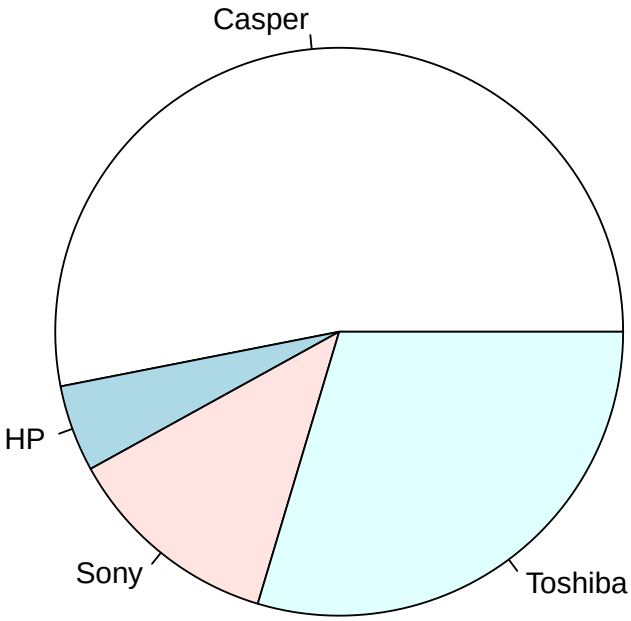
Substitution rates

```
subst_df <- market_df %>%
  select(id, test_data) %>%
  unnest(test_data) %>%
  select(id, Brand, RANK) %>%
  filter(RANK<=2) %>%
  pivot_wider(names_from = "RANK", values_from = "Brand") %>%
  count(`1`,`2`) %>%
  rename(Brand = `1`, Substitute = `2`) %>%
  nest(Subs = -Brand)

oldpar <- par(mar = c(0,0,1,0))

for (i in seq(nrow(subst_df))){
  subst_df$Subs[[i]] %>%
      column_to_rownames(var = "Substitute") %>%
      as.matrix %>%
      drop() %>%
      pie(main=paste("Substitutes of", subst_df$Brand[i] ))
}
par(mar = oldpar$mar)
```









Market simulation

All possible Casper configurations

```
config_df <- test_df %>%
  select(Processor, Speed, Memory, HardDisk, Screen, Price) %>%
  expand(Processor, Speed, Memory, HardDisk, Screen, Price) %>% # expand to all permutations of the val
  mutate(Brand = "Casper") %>% # our company is Casper
  relocate(Brand) %>%
```

```
  anti_join(test_df, by = c("Brand", "Processor", "Speed", "Memory", "HardDisk", "Screen", "Price")) %>%
  mutate(config = as.character(seq(n()))) %>%
  relocate(config)
```

Predicted rank utilities of all configurations

```
market_config_df <- market_df %>%
  mutate(config_pred = map2(model, id, ~{
    pred <- predict(.x, config_df) %>%
      `*`(-1)
    config_df %>%
      select(config) %>%
      mutate(PREDraw = pred)
  })) %>%
  select(id, test_pred, config_pred) %>%
  unnest(test_pred) %>%
  select(-RANK, -PREDrank) %>%
  pivot_wider(names_from = Alternative, values_from = "PREDraw") %>%
  unnest(config_pred) %>%
  rename(new = PREDraw)

market_config_df %>%
  head(10)
```

```
## # A tibble: 10 x 9
##    id          config   new   `1`   `2`   `3`   `4`   `5`   `6`
##    <chr>       <chr>  <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
##  1 y14s208s001 1       15.5    12     8   7.5    11     6  12.5
##  2 y14s208s001 2       10.5    12     8   7.5    11     6  12.5
##  3 y14s208s001 3       11.5    12     8   7.5    11     6  12.5
##  4 y14s208s001 4       11.5    12     8   7.5    11     6  12.5
##  5 y14s208s001 5       16.5    12     8   7.5    11     6  12.5
##  6 y14s208s001 6       11.5    12     8   7.5    11     6  12.5
##  7 y14s208s001 7       12.5    12     8   7.5    11     6  12.5
##  8 y14s208s001 8       12.5    12     8   7.5    11     6  12.5
##  9 y14s208s001 9       16.5    12     8   7.5    11     6  12.5
## 10 y14s208s001 10      11.5    12     8   7.5    11     6  12.5
```

New markets obtained by adding each config to the existing hepsiburada market

```
new_markets_df <- market_config_df %>%
  pivot_longer(cols = c(-id, -config), names_to = "Alternative", values_to = "PREDraw") %>%
  group_by(id, config) %>%
  mutate(PREDrank = rank(PREDraw, ties.method = "first")) %>%
  ungroup() %>%
  left_join(mutate_all(select(test_hande_df, Alternative, Brand), as.character), by = "Alternative") %>%
  replace_na(list(Brand = "Casper")) %>%
  nest(new_market = -config)

new_markets_df  %>%
  head(10)
```

```
## # A tibble: 10 x 2
##    config new_market
##    <chr>  <list>
##  1 1      <tibble [1,295 x 5]>
```

```
##  2 2      <tibble [1,295 x 5]>
##  3 3      <tibble [1,295 x 5]>
##  4 4      <tibble [1,295 x 5]>
##  5 5      <tibble [1,295 x 5]>
##  6 6      <tibble [1,295 x 5]>
##  7 7      <tibble [1,295 x 5]>
##  8 8      <tibble [1,295 x 5]>
##  9 9      <tibble [1,295 x 5]>
## 10 10     <tibble [1,295 x 5]>
```

```
new_markets_df$new_market[[1]] %>%
  filter(PREDrank == 1) %>%
  count(Brand) %>%
  mutate(Perc = 100*n/sum(n))
```

```
## # A tibble: 4 x 3
##   Brand       n  Perc
##   <chr>   <int> <dbl>
## 1 Casper     81 43.8
## 2 HP         11  5.95
## 3 Sony       22 11.9
## 4 Toshiba    71 38.4
```

Market shares after each new configuration is introduced to the existing hepsiburada market

```
new_markets_stat_df <- new_markets_df %>%
  mutate(market_shares = map(new_market, ~{
    .x %>%
      filter(PREDrank == 1) %>%
      count(Brand) %>%
      mutate(Perc = 100*n/sum(n))
  }),
  Casper_mshare = map_dbl(market_shares, ~{.x %>% filter(Brand == "Casper") %>% pull("Perc")})) %>%
  left_join(config_df, by = "config") %>%
  mutate_at(vars(Memory, Price), ~{parse_number(as.character(.), locale = locale(grouping_mark = "."))})
  arrange(desc(Casper_mshare), Price)

new_markets_stat_df %>%
  select(-new_market, -market_shares) %>%
  filter(Casper_mshare >= mshares["Casper"])
```

```
## # A tibble: 78 x 9
##    config Casper_mshare Brand  Processor     Speed  Memory HardDisk Screen Price
##    <chr>          <dbl> <chr>  <fct>         <fct>   <dbl> <fct>    <fct>  <dbl>
##  1 228             59.5 Casper Intel Core i7 2.6 G~     16 1.000 GB 17.3 ~  6000
##  2 224             58.4 Casper Intel Core i7 2.6 G~     16 1.000 GB 15.6 ~  6000
##  3 230             57.8 Casper Intel Core i7 2.6 G~     16 1.000 GB 17.3 ~  8500
##  4 232             57.3 Casper Intel Core i7 2.6 G~     16 750 GB   15.6 ~  6000
##  5 225             57.3 Casper Intel Core i7 2.6 G~     16 1.000 GB 15.6 ~  7000
##  6 229             57.3 Casper Intel Core i7 2.6 G~     16 1.000 GB 17.3 ~  7000
##  7 236             56.2 Casper Intel Core i7 2.6 G~     16 750 GB   17.3 ~  6000
##  8 226             56.2 Casper Intel Core i7 2.6 G~     16 1.000 GB 15.6 ~  8500
##  9 237             55.1 Casper Intel Core i7 2.6 G~     16 750 GB   17.3 ~  7000
## 10 238             55.1 Casper Intel Core i7 2.6 G~     16 750 GB   17.3 ~  8500
## # ... with 68 more rows
```

```
new_markets_stat_df %>%
  filter(Casper_mshare >= mshares["Casper"]) %>%
  select(-new_market, -market_shares) %>%
  ggplot(aes(Price, Casper_mshare)) +
  geom_hline(yintercept = mshares["Casper"], size=2, col = "yellow") +
  geom_jitter(aes(col = Screen), width=200, height=0, alpha = 0.8, size=3) +
  scale_x_continuous(labels = scales::comma) +
  labs(title = paste("Current Casper market share is", round(mshares["Casper"]), "% (yellow horizontal l
       y = "Predicted Casper market share (%)") +
  facet_grid(Memory~HardDisk, as.table = FALSE, labeller = label_both) +
  theme(strip.text.y = element_text(angle=0))
```