Questions vs directives

Question
“Does treatment duration have an effect on survival?”

Directive
“Make a figure of survival probability as a function of treatment duration.”

Questions end in a question mark!
Conceptual vs procedural questions

Conceptual question
“Does treatment duration have an effect on survival?”

Procedural question
“What is the difference in mean survival between a treatment duration of 1 month and of 2 months?”

Conceptual questions do not prompt a specific analysis procedure!
Working with tidy data in R: dplyr

Fundamental actions on data tables:

• choose rows — filter() 
• choose columns — select() 
• make new columns — mutate() 
• arrange rows — arrange() 
• calculate summary statistics — summarize() 
• work on groups of data — group_by()
We can combine these verbs using the pipe operator: `%>%`

Standard R:

```r
> mean(iris$Sepal.Length)
[1] 5.843333
```

With pipe:

```r
> iris$Sepal.Length %>% mean()
[1] 5.843333
```
We can combine these verbs using the pipe operator: `%>%`

Standard R:

```
> head(iris)
             Sepal.Length  Sepal.Width  Petal.Length  Petal.Width Species
1           5.1         3.5          1.4         0.2     setosa
2           4.9         3.0          1.4         0.2     setosa
3           4.7         3.2          1.3         0.2     setosa
4           4.6         3.1          1.5         0.2     setosa
5           5.0         3.6          1.4         0.2     setosa
6           5.4         3.9          1.7         0.4     setosa
```
We can combine these verbs using the pipe operator: `%%`

With pipe:

```r
> iris %>% head()

  Sepal.Length Sepal.Width Petal.Length Petal.Width Species
  1         5.1       3.5        1.4      0.2  setosa
  2         4.9       3.0        1.4      0.2  setosa
  3         4.7       3.2        1.3      0.2  setosa
  4         4.6       3.1        1.5      0.2  setosa
  5         5.0       3.6        1.4      0.2  setosa
  6         5.4       3.9        1.7      0.4  setosa
```
Left and right assignment: <- and ->

Left assignment:
> x <- 5
> x
[1] 5

Right assignment:
> 6 -> x
> x
[1] 6
Combining pipe and right assignment

These three lines do all the same thing:

```r
> mean_length <- mean(iris$Sepal.Length)
> mean_length <- iris$Sepal.Length %>% mean()
> iris$Sepal.Length %>% mean() -> mean_length
> mean_length
[1] 5.843333
```
Pipe example 1: count how many herbivores of different orders there are in msleep
Pipe example 1: count how many herbivores of different orders there are in \texttt{msleep}

\begin{verbatim}
msleep %>%
  filter(vore == "herbi")
\end{verbatim}
Pipe example 1: count how many herbivores of different orders there are in msleep

```r
msleep %>%
  filter(vore == "herbi") %>%
  group_by(order)
```
Pipe example 1: count how many herbivores of different orders there are in `msleep`

```r
msleep %>%
  filter(vore == "herbi") %>%
  group_by(order) %>%
  summarize(count = n())
```
Pipe example 1: count how many herbivores of different orders there are in `msleep`

```r
msleep %>%
  filter(vore == "herbi") %>%
  group_by(order) %>%
  summarize(count = n()) %>%
  arrange(desc(count))
```
Pipe example 1: count how many herbivores of different orders there are in `msleep`

```r
msleep %>%
  filter(vore == "herbi") %>%
  group_by(order) %>%
  summarize(count = n()) %>%
  arrange(desc(count))
```

<table>
<thead>
<tr>
<th>order</th>
<th>count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rodentia</td>
<td>16</td>
</tr>
<tr>
<td>Artiodactyla</td>
<td>5</td>
</tr>
<tr>
<td>Perissodactyla</td>
<td>3</td>
</tr>
<tr>
<td>Hyracoidea</td>
<td>2</td>
</tr>
<tr>
<td>Proboscidea</td>
<td>2</td>
</tr>
<tr>
<td>Diprotodontia</td>
<td>1</td>
</tr>
<tr>
<td>Lagomorpha</td>
<td>1</td>
</tr>
<tr>
<td>Pilosa</td>
<td>1</td>
</tr>
<tr>
<td>Primates</td>
<td>1</td>
</tr>
</tbody>
</table>
Pipe example 2: What is total day time for each animal in msleep?
Pipe example 2: What is total day time for each animal in msleep?

```r
msleep %>%
  mutate(total_day_time = awake + sleep_total)
```
Pipe example 2: What is total day time for each animal in msleep?

```r
msleep %>%
  mutate(total_day_time = awake + sleep_total) %>%
  select(name, total_day_time)
```
Pipe example 2: What is total day time for each animal in msleep?

```r
msleep %>%
  mutate(total_day_time = awake + sleep_total) %>%
  select(name, total_day_time)
```

<table>
<thead>
<tr>
<th>name</th>
<th>total_day_time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cheetah</td>
<td>24.00</td>
</tr>
<tr>
<td>Owl monkey</td>
<td>24.00</td>
</tr>
<tr>
<td>Mountain beaver</td>
<td>24.00</td>
</tr>
<tr>
<td>Greater short-tailed shrew</td>
<td>24.00</td>
</tr>
<tr>
<td>Cow</td>
<td>24.00</td>
</tr>
<tr>
<td>Three-toed sloth</td>
<td>24.00</td>
</tr>
<tr>
<td>Northern fur seal</td>
<td>24.00</td>
</tr>
<tr>
<td>Vesper mouse</td>
<td>24.00</td>
</tr>
<tr>
<td>Dog</td>
<td>24.00</td>
</tr>
<tr>
<td>Roe deer</td>
<td>24.00</td>
</tr>
</tbody>
</table>
Pipe example 3: What is the median awake time of different orders in `msleep`?
Pipe example 3: What is the median awake time of different orders in msleep?

```r
msleep %>%
  group_by(order)
```
Pipe example 3: What is the median awake time of different orders in `msleep`?

```r
msleep %>%
  group_by(order) %>%
  summarize(med_awake = median(awake))
```
Pipe example 3: What is the median awake time of different orders in `msleep`?

```r
msleep %>%
  group_by(order) %>%
  summarize(med_awake = median(awake)) %>%
  arrange(med_awake)
```
Pipe example 3: What is the median awake time of different orders in msleep?

msleep %>%
  group_by(order) %>%
  summarize(med_awake = median(awake)) %>%
  arrange(med_awake)

<table>
<thead>
<tr>
<th>order</th>
<th>med_awake</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chiroptera</td>
<td>4.20</td>
</tr>
<tr>
<td>Didelphimorphia</td>
<td>5.30</td>
</tr>
<tr>
<td>Cingulata</td>
<td>6.25</td>
</tr>
<tr>
<td>Afrosoricida</td>
<td>8.40</td>
</tr>
<tr>
<td>Pilosa</td>
<td>9.60</td>
</tr>
<tr>
<td>Rodentia</td>
<td>11.10</td>
</tr>
<tr>
<td>Diprotodontia</td>
<td>11.60</td>
</tr>
<tr>
<td>Soricomorpha</td>
<td>13.70</td>
</tr>
<tr>
<td>Carnivora</td>
<td>13.75</td>
</tr>
<tr>
<td>Erinaceomorpha</td>
<td>13.80</td>
</tr>
</tbody>
</table>