In [5]:
```
import pandas as pd
```

In [6]:
```
import plotnine as p9
```

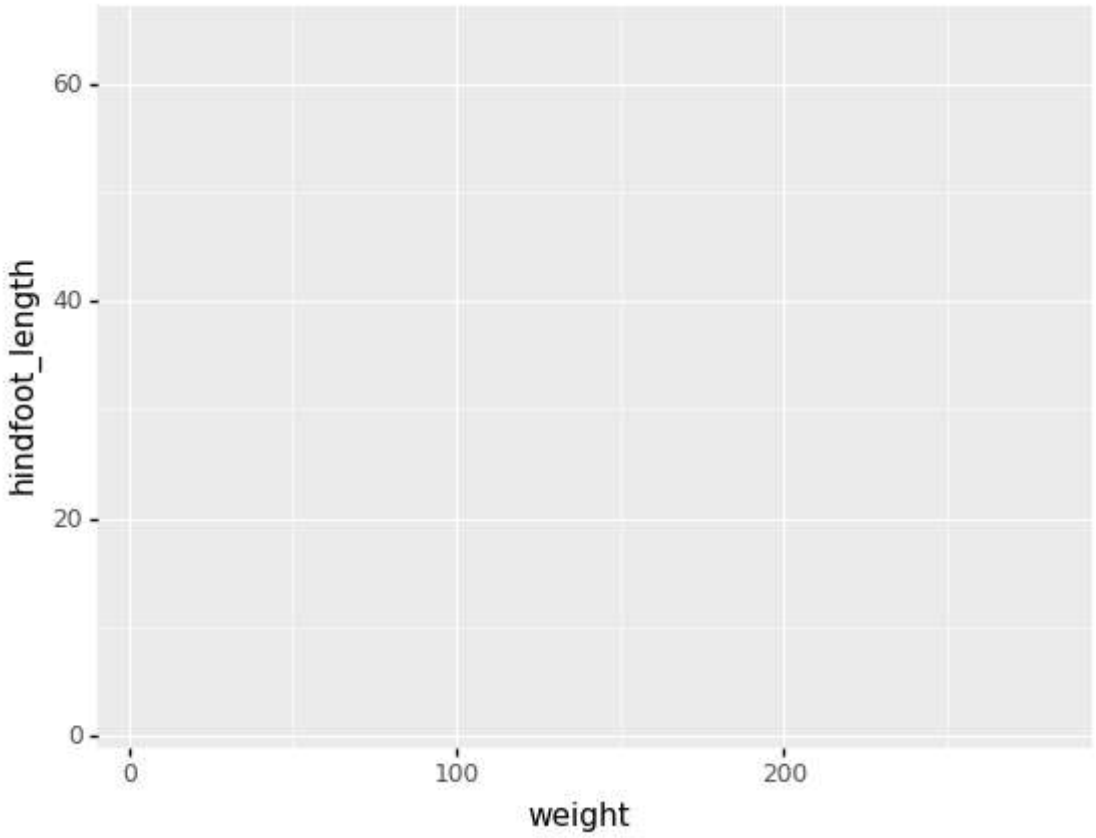
In [7]:
```
surveys_complete = pd.read_csv('surveys.csv')
```

In [8]:
```
surveys_complete = surveys_complete.dropna()
```

In [11]:
```
(p9.ggplot(data=surveys_complete))
```

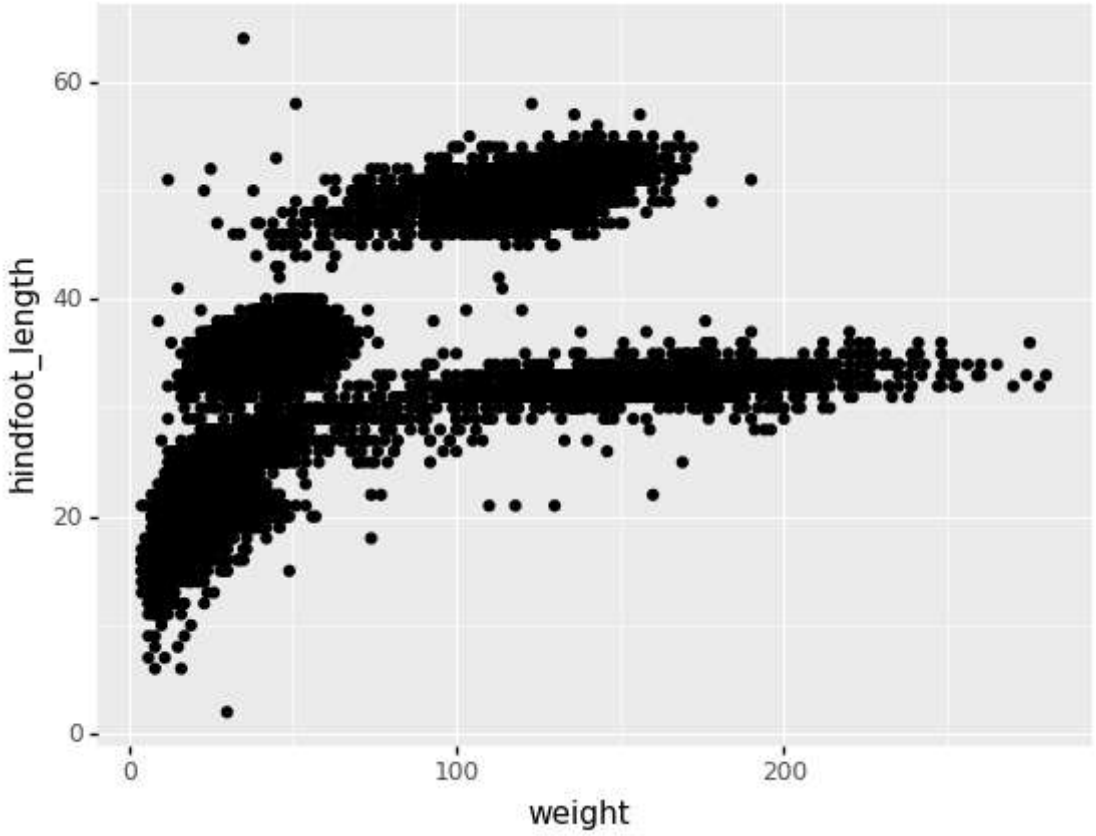

Out[11]: <ggplot: (128965688240)>

In [12]:
```
(p9.ggplot(data=surveys_complete,
           mapping=p9.aes(x='weight', y='hindfoot_length')))
```



Out[12]: <ggplot: (128966079678)>

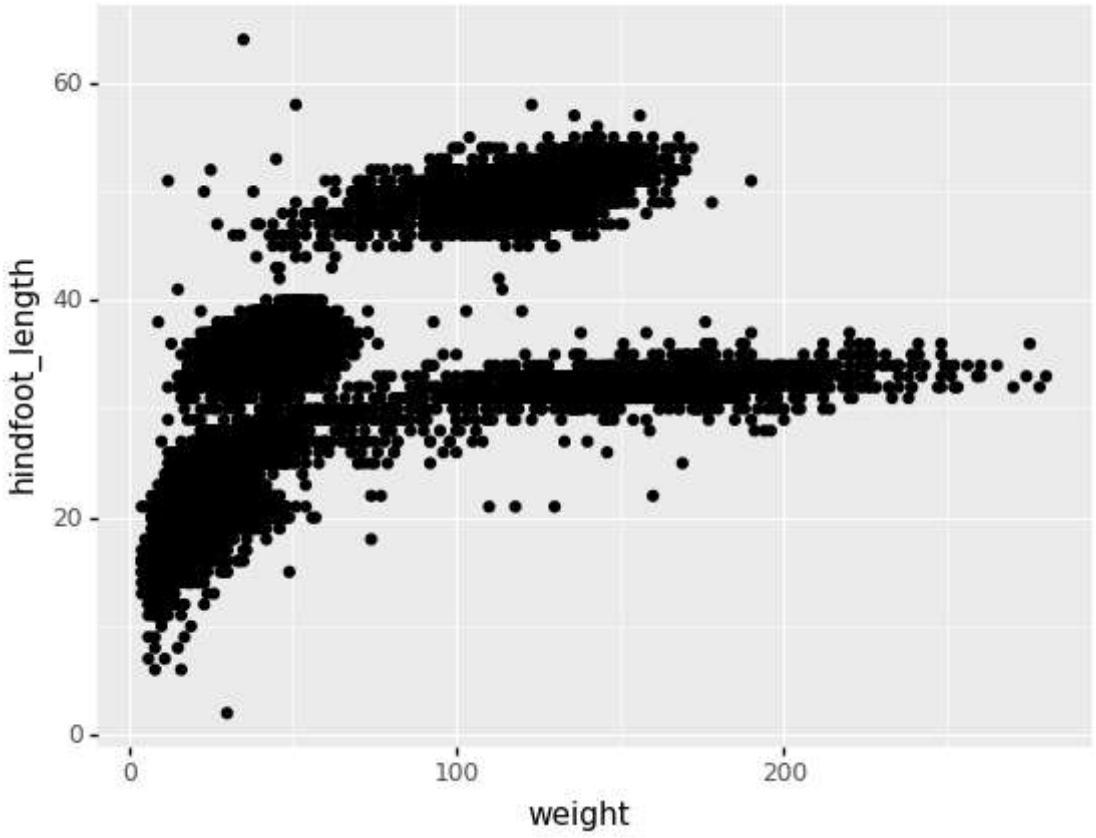
In [13]:
```
(p9.ggplot(data=surveys_complete,
           mapping=p9.aes(x='weight', y='hindfoot_length'))
    + p9.geom_point()
)
```



Out[13]: <ggplot: (128966132849)>

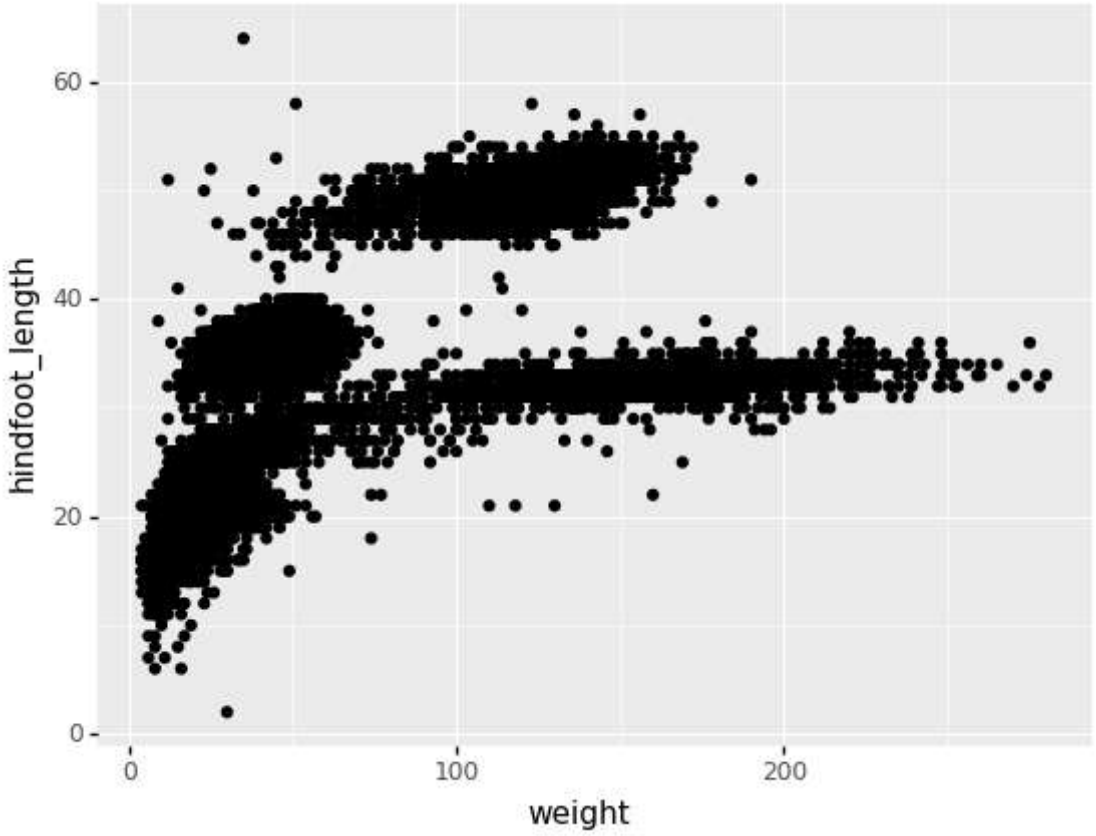
In [14]:
```
# Create
surveys_plot = p9.ggplot(data=surveys_complete,
                         mapping=p9.aes(x='weight', y='hindfoot_length'))

# Draw the plot
surveys_plot + p9.geom_point()
```



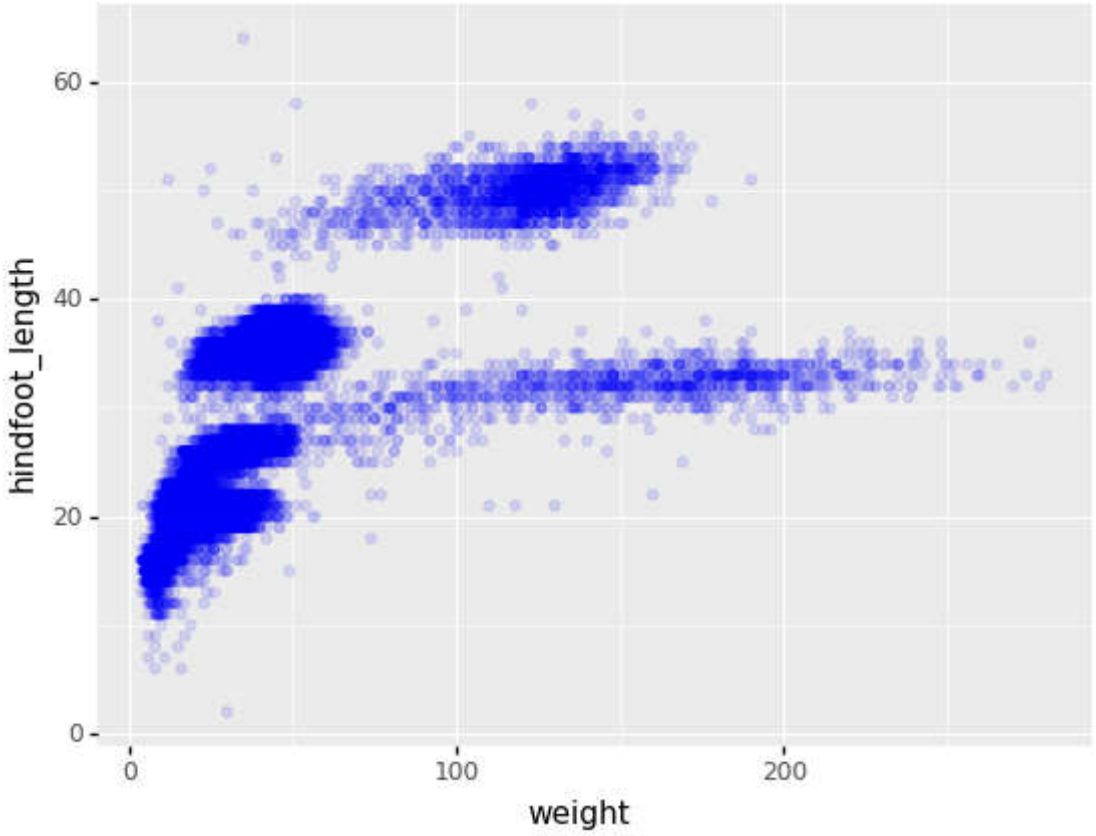
Out[14]: <ggplot: (128971261476)>

In [15]:
```
(p9.ggplot(data=surveys_complete,
           mapping=p9.aes(x='weight', y='hindfoot_length'))
    + p9.geom_point()
)
```



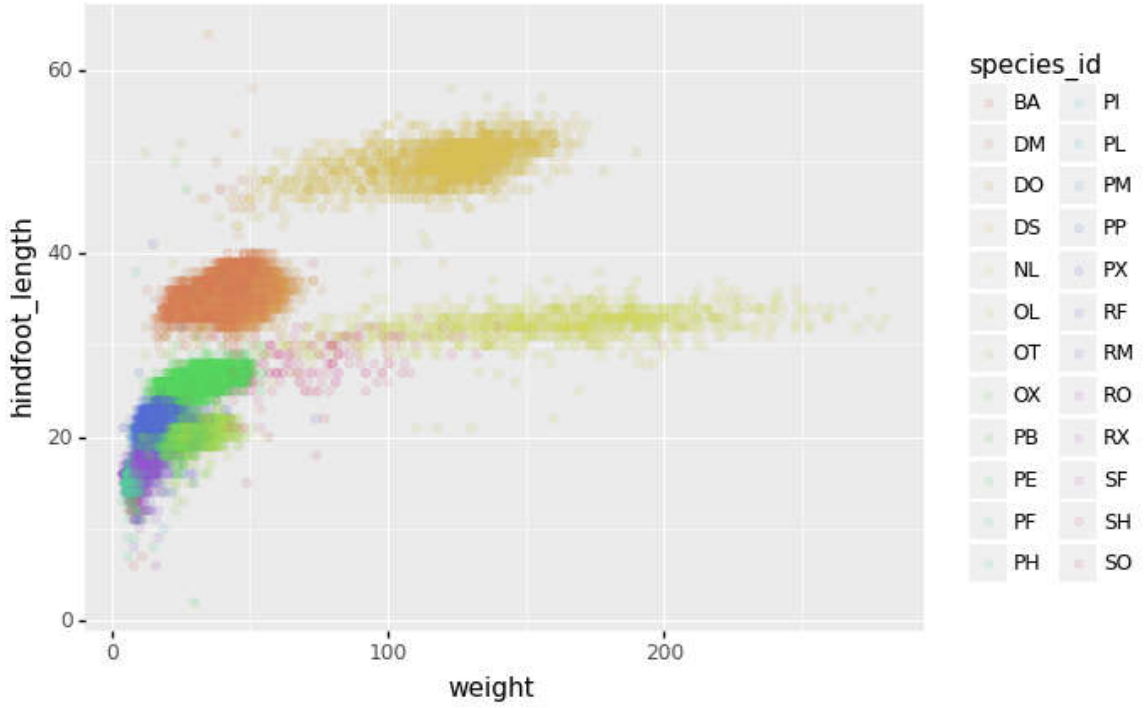
Out[15]: <ggplot: (-9223371907888541475)>

In [16]:
```
(p9.ggplot(data=surveys_complete,
           mapping=p9.aes(x='weight', y='hindfoot_length'))
    + p9.geom_point(alpha=0.1, color='blue')
)
```



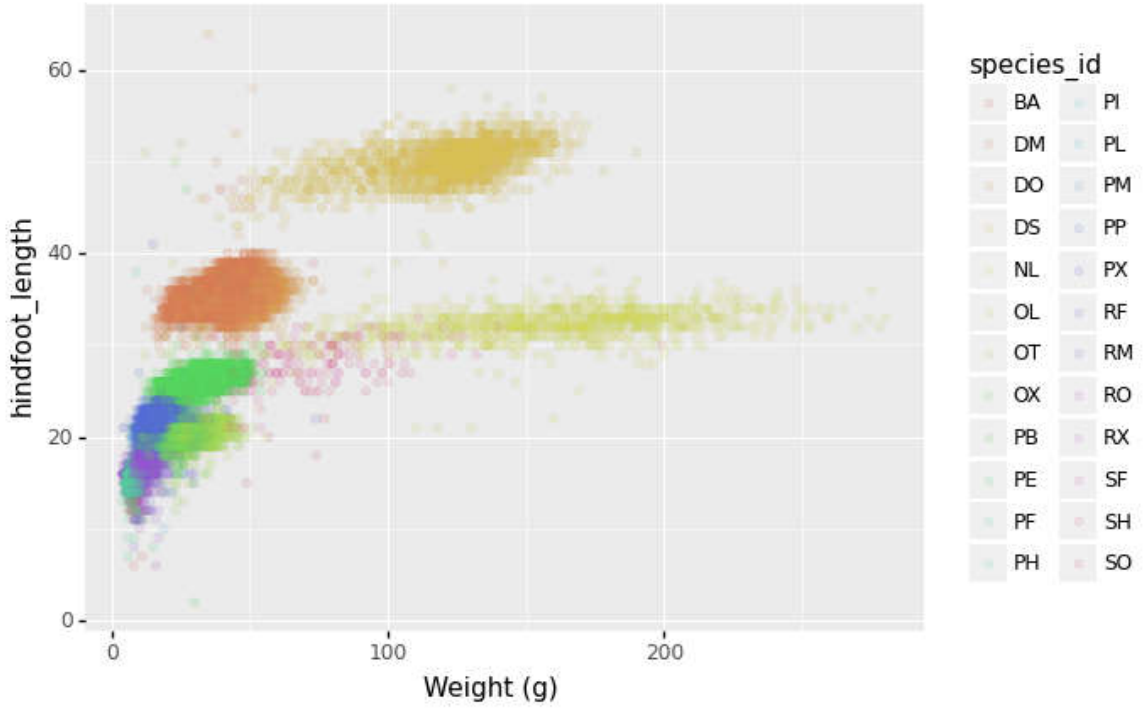
Out[16]: <ggplot: (-9223371907888307119)>

In [17]:
```
(p9.ggplot(data=surveys_complete,
           mapping=p9.aes(x='weight',
                          y='hindfoot_length',
                          color='species_id'))
    + p9.geom_point(alpha=0.1)
)
```



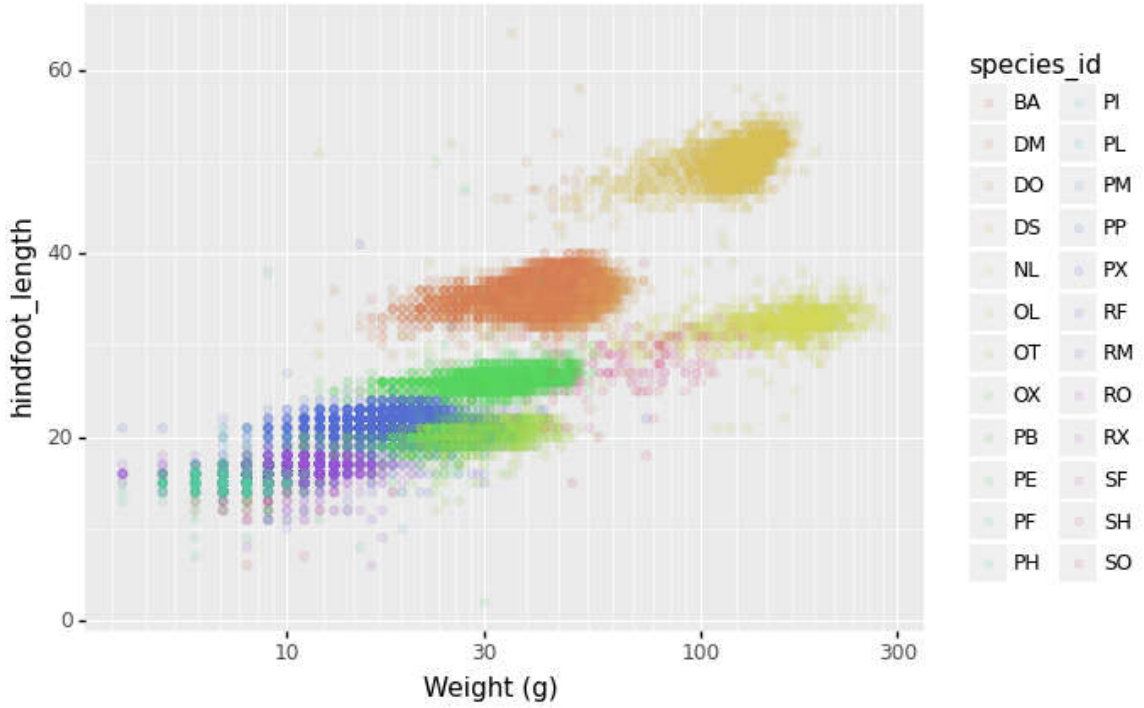
Out[17]: <ggplot: (-9223371907882368459)>

In [18]:
```
(p9.ggplot(data=surveys_complete,
           mapping=p9.aes(x='weight', y='hindfoot_length', color='species_id'))
    + p9.geom_point(alpha=0.1)
    + p9.xlab("Weight (g)")
)
```



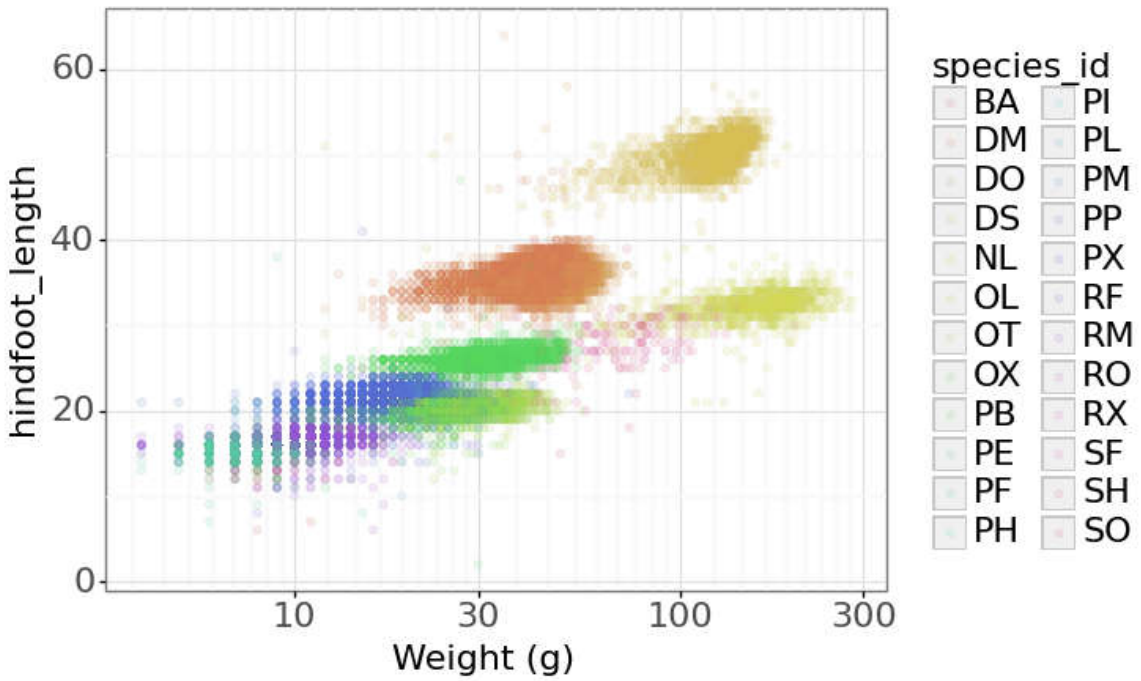
Out[18]: <ggplot: (-9223371907888185996)>

In [19]:
```
(p9.ggplot(data=surveys_complete,
           mapping=p9.aes(x='weight', y='hindfoot_length', color='species_id'))
    + p9.geom_point(alpha=0.1)
    + p9.xlab("Weight (g)")
    + p9.scale_x_log10()
)
```



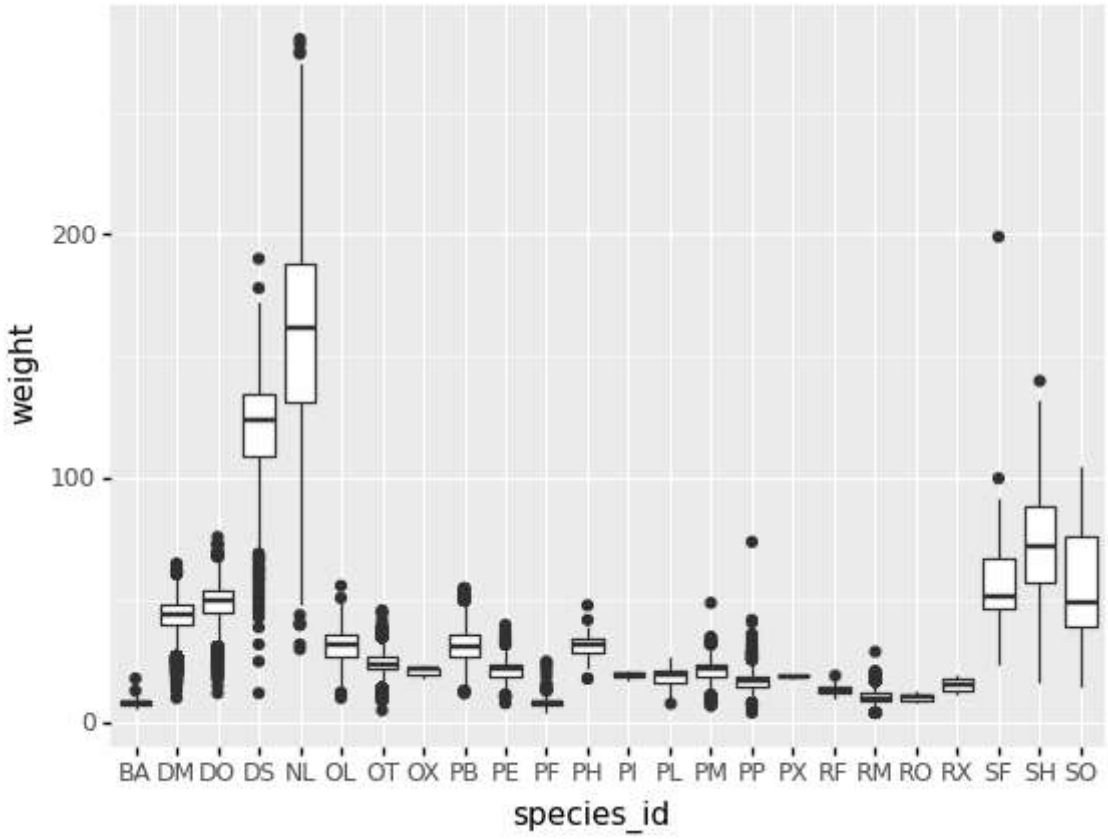
Out[19]: <ggplot: (-9223371907883002666)>

In [20]:
```
(p9.ggplot(data=surveys_complete,
           mapping=p9.aes(x='weight', y='hindfoot_length', color='species_id'))
    + p9.geom_point(alpha=0.1)
    + p9.xlab("Weight (g)")
    + p9.scale_x_log10()
    + p9.theme_bw()
    + p9.theme(text=p9.element_text(size=16))
)
```



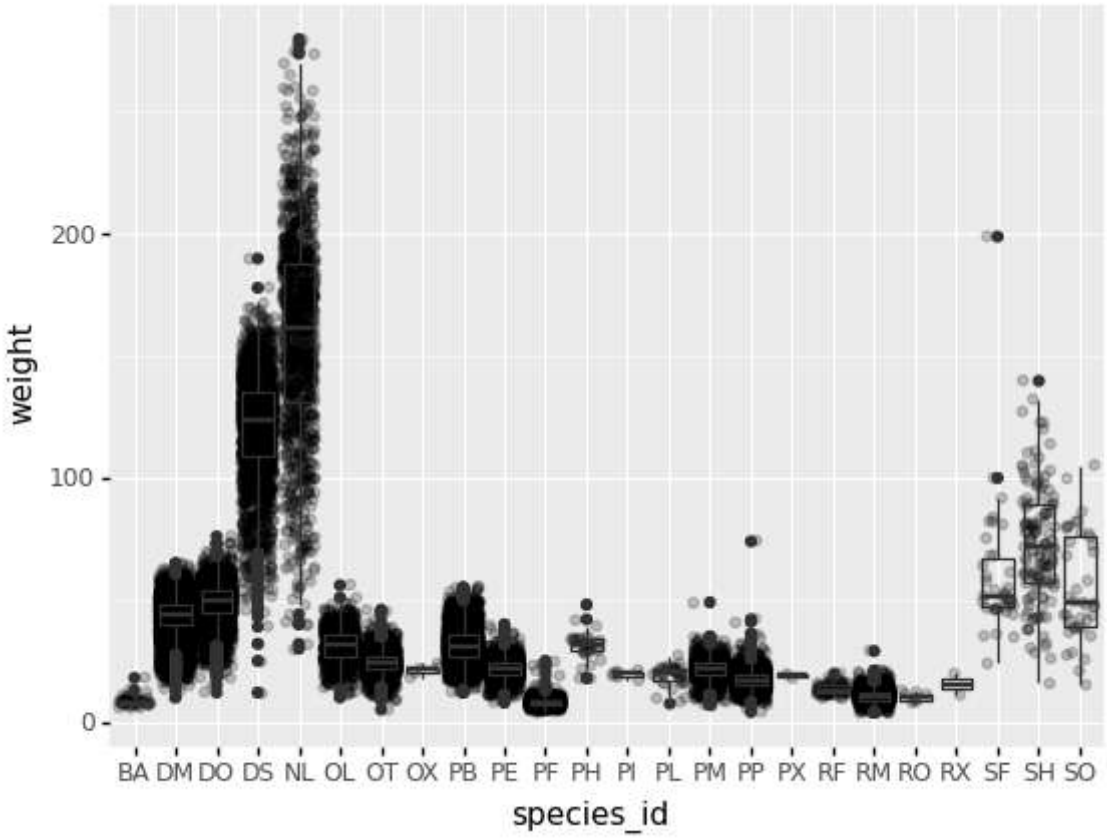
Out[20]: <ggplot: (128972225768)>

In [21]:
```
(p9.ggplot(data=surveys_complete,
           mapping=p9.aes(x='species_id',
                          y='weight'))
    + p9.geom_boxplot()
)
```



Out[21]: <ggplot: (-9223371907882993948)>

In [22]:
```
(p9.ggplot(data=surveys_complete,
           mapping=p9.aes(x='species_id',
                          y='weight'))
    + p9.geom_jitter(alpha=0.2)
    + p9.geom_boxplot(alpha=0.)
)
```



Out[22]: <ggplot: (128972222472)>

In [23]:
```
yearly_counts = surveys_complete.groupby(['year', 'species_id'])['species_id'].count()
yearly_counts
```

Out[23]:
```
year  species_id
1977  DM            181
      DO             12
      DS             29
      OL              1
      OX              2
                   ...
2002  PP            375
      RM             20
      RO              7
      SF              5
      SH              9
Name: species_id, Length: 332, dtype: int64
```

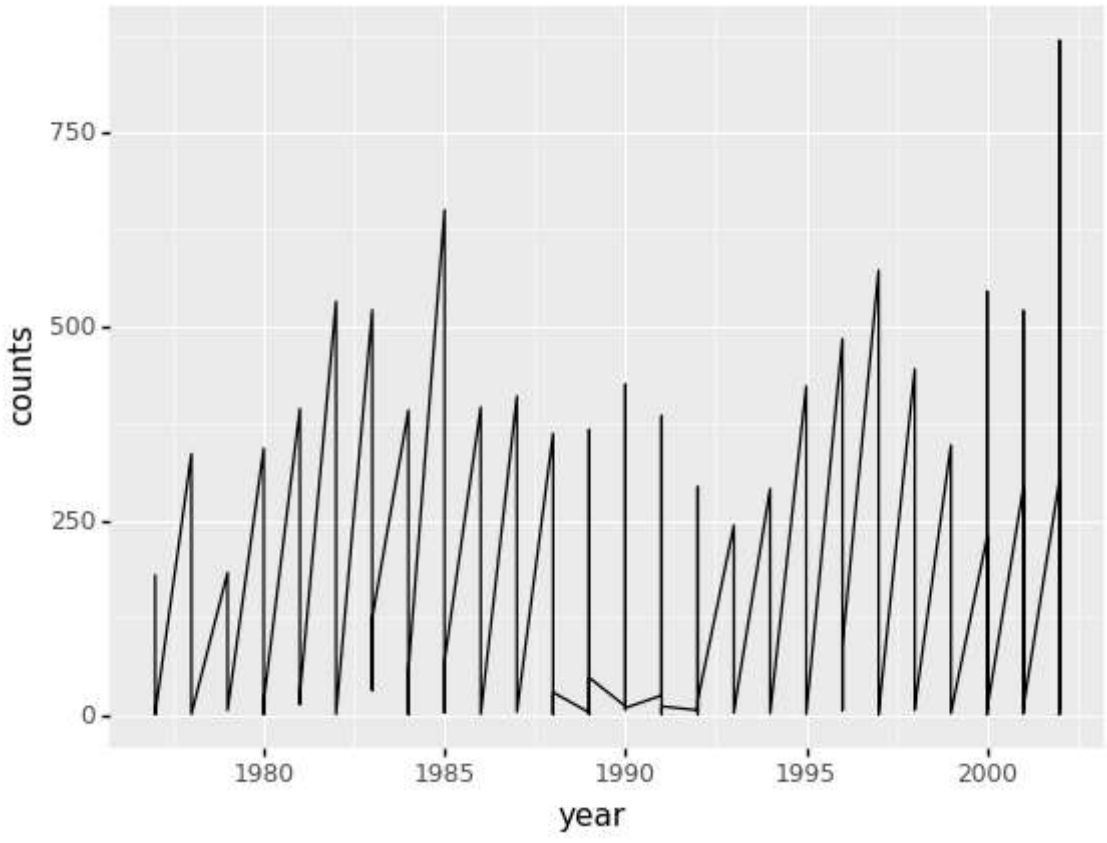
In [24]:
```
yearly_counts = yearly_counts.reset_index(name='counts')
yearly_counts
```

Out[24]:

	year	species_id	counts
0	1977	DM	181
1	1977	DO	12
2	1977	DS	29
3	1977	OL	1
4	1977	OX	2
...	...	...	...
327	2002	PP	375
328	2002	RM	20
329	2002	RO	7
330	2002	SF	5
331	2002	SH	9

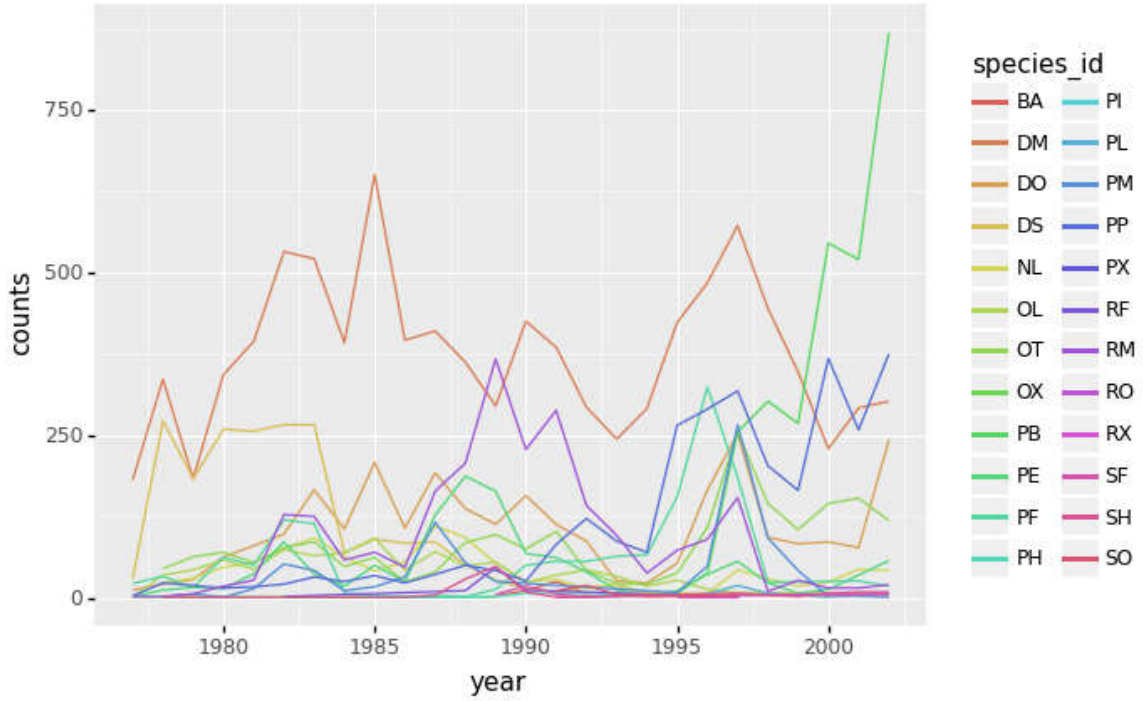
332 rows × 3 columns

In [25]:
```
(p9.ggplot(data=yearly_counts,
           mapping=p9.aes(x='year',
                          y='counts'))
    + p9.geom_line()
)
```



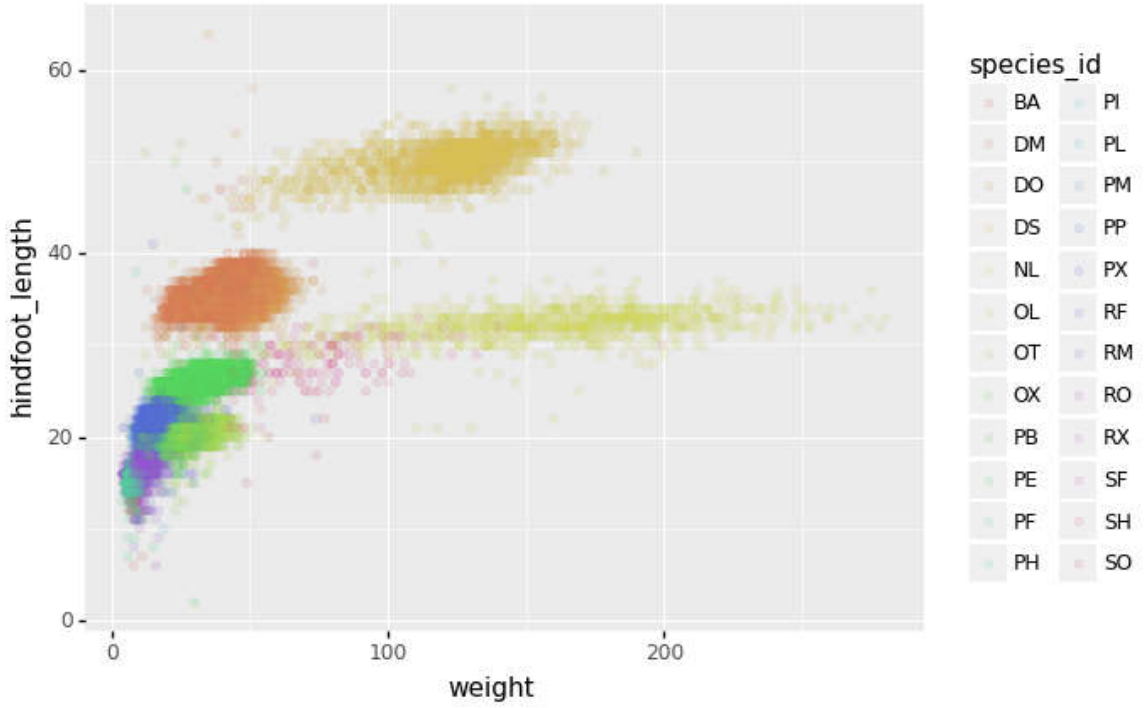
Out[25]: <ggplot: (128938421256)>

In [26]:
```
(p9.ggplot(data=yearly_counts,
           mapping=p9.aes(x='year',
                          y='counts',
                          color='species_id'))
    + p9.geom_line()
)
```



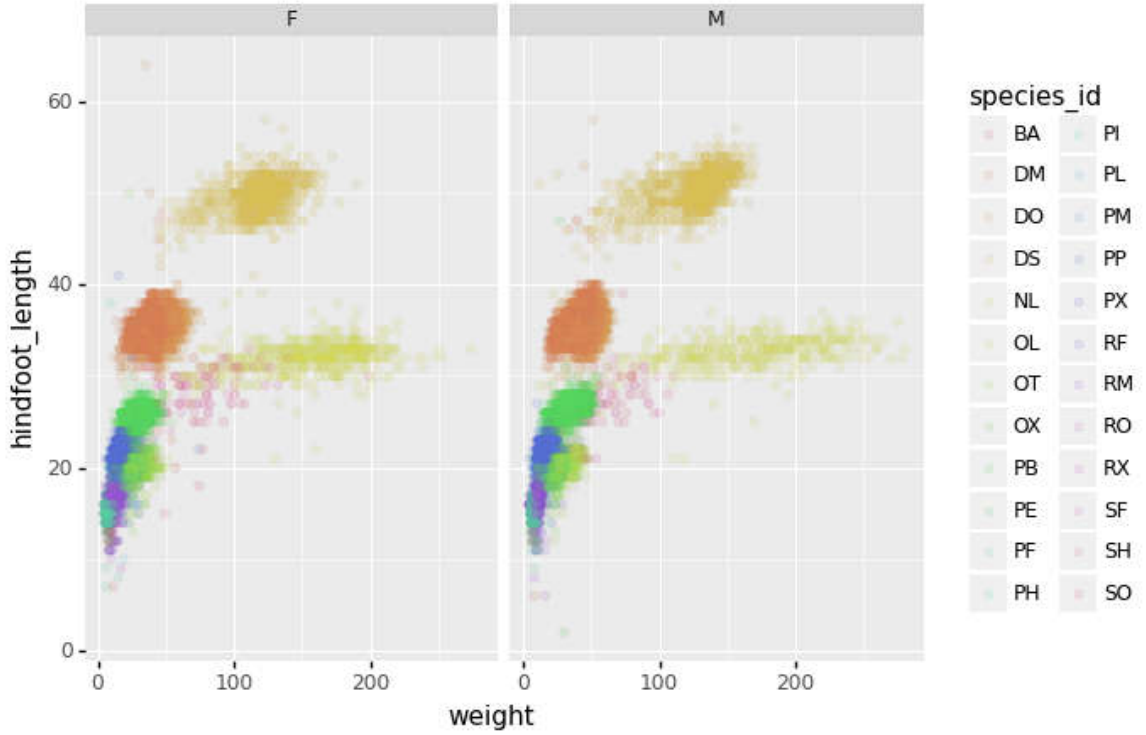
Out[26]: <ggplot: (128965694710)>

In [28]:
```
(p9.ggplot(data=surveys_complete,
           mapping=p9.aes(x='weight',
                          y='hindfoot_length',
                          color='species_id'))
    + p9.geom_point(alpha=0.1)
)
```



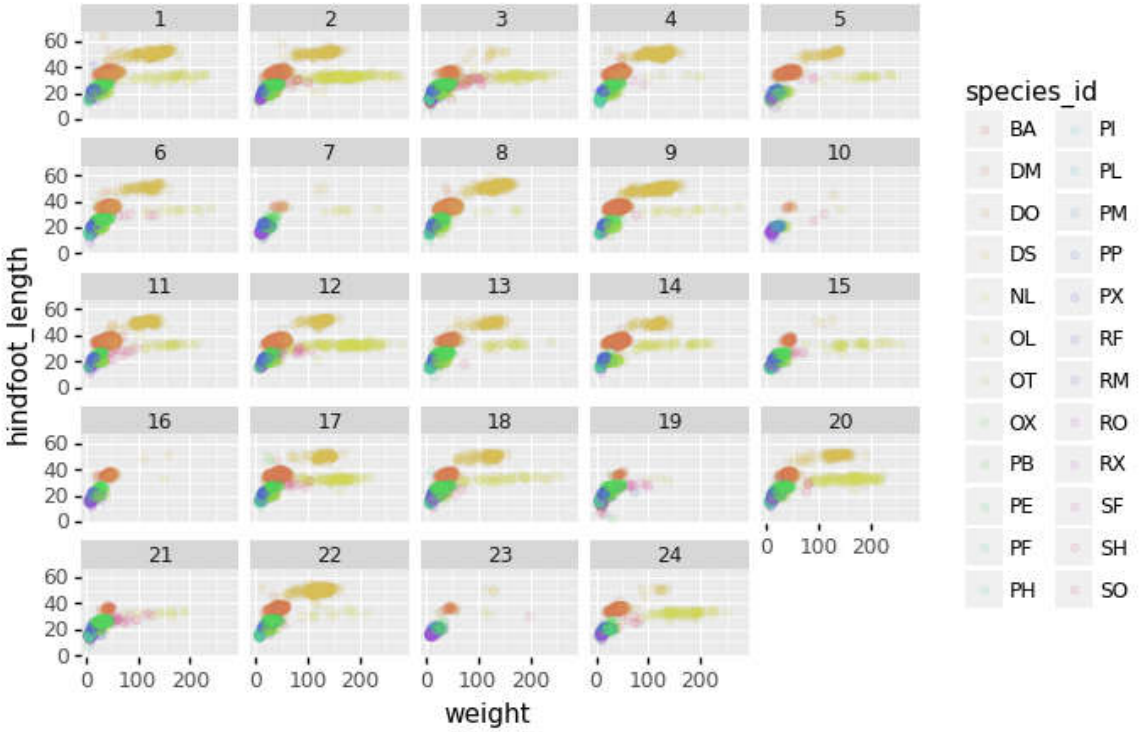
Out[28]: <ggplot: (-9223371907886949572)>

In [29]:
```
(p9.ggplot(data=surveys_complete,
           mapping=p9.aes(x='weight',
                          y='hindfoot_length',
                          color='species_id'))
    + p9.geom_point(alpha=0.1)
    + p9.facet_wrap("sex")
)
```



Out[29]: <ggplot: (128971144175)>

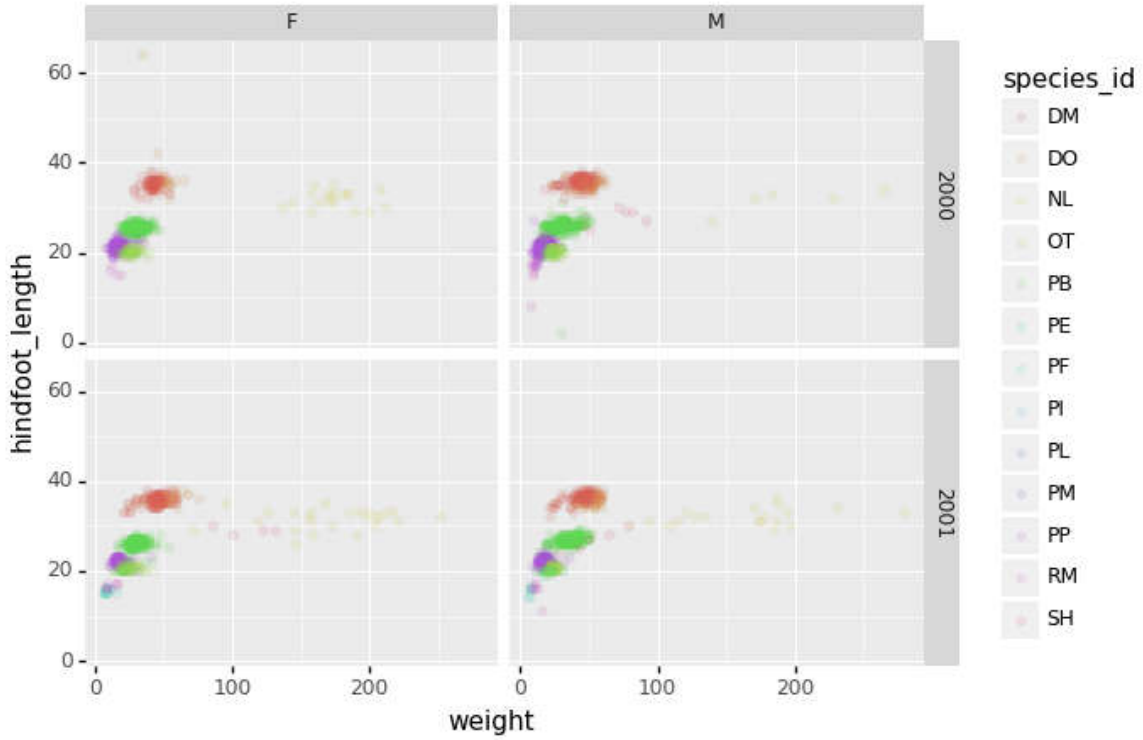
In [30]:
```
(p9.ggplot(data=surveys_complete,
           mapping=p9.aes(x='weight',
                          y='hindfoot_length',
                          color='species_id'))
    + p9.geom_point(alpha=0.1)
    + p9.facet_wrap("plot_id")
)
```



Out[30]: <ggplot: (-9223371907883687073)>

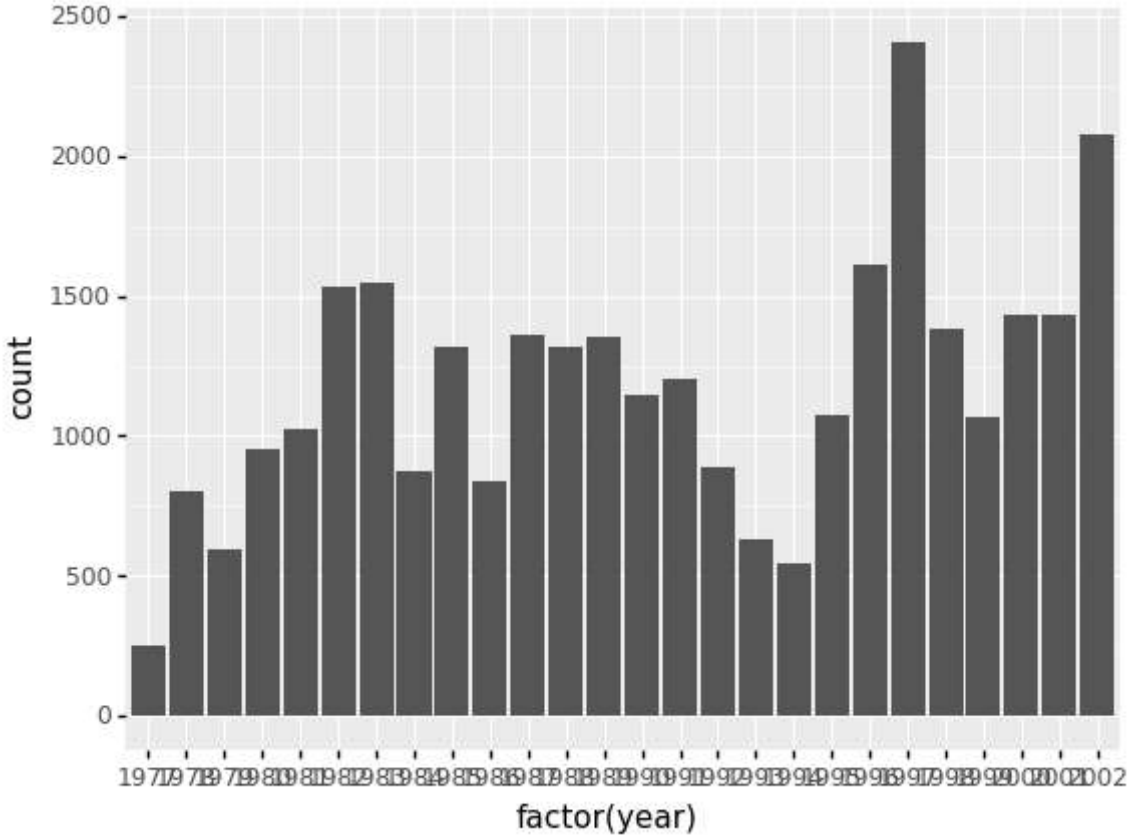
In [31]:
```
# only select the years of interest
survey_2000 = surveys_complete[surveys_complete["year"].isin([2000, 2001])]

(p9.ggplot(data=survey_2000,
           mapping=p9.aes(x='weight',
                          y='hindfoot_length',
                          color='species_id'))
    + p9.geom_point(alpha=0.1)
    + p9.facet_grid("year ~ sex")
)
```



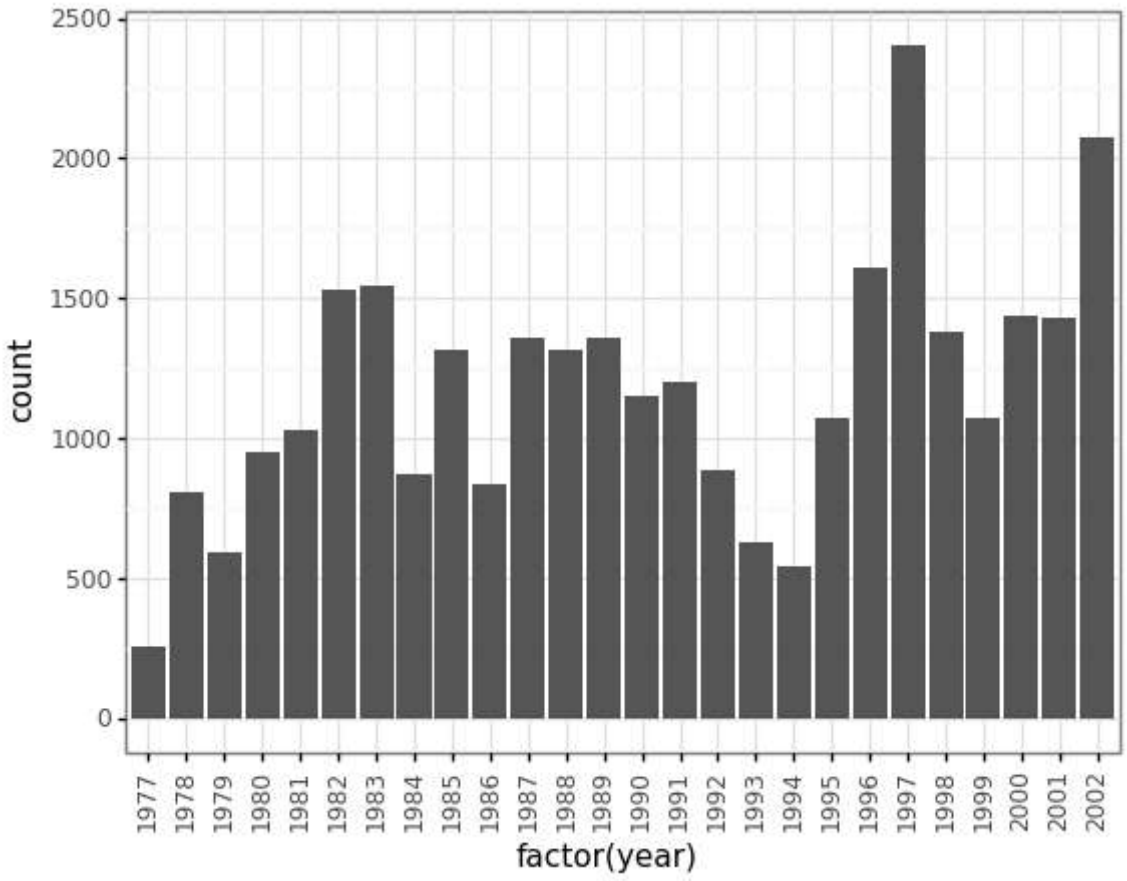
Out[31]: <ggplot: (128971146886)>

In [32]:
```
(p9.ggplot(data=surveys_complete,
           mapping=p9.aes(x='factor(year)'))
    + p9.geom_bar()
)
```



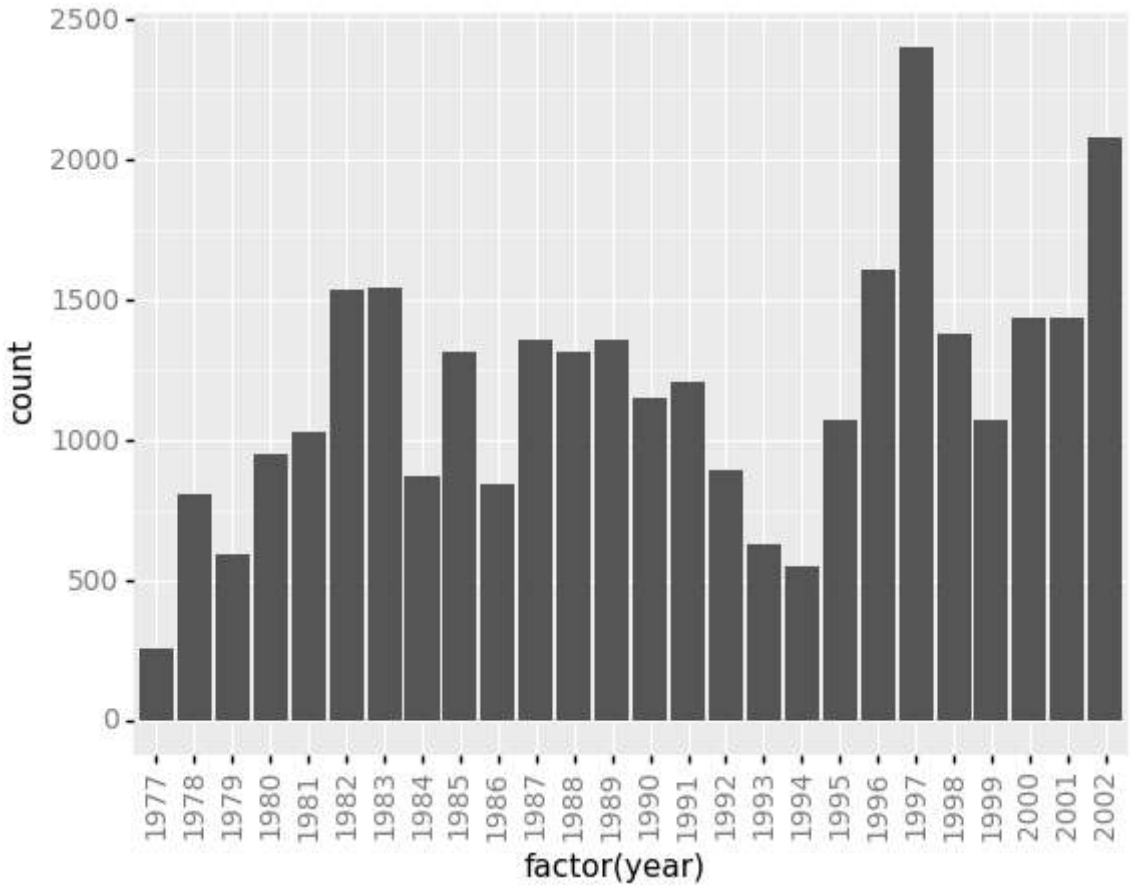
Out[32]: <ggplot: (-9223371907888701173)>

In [33]:
```
(p9.ggplot(data=surveys_complete,
           mapping=p9.aes(x='factor(year)'))
    + p9.geom_bar()
    + p9.theme_bw()
    + p9.theme(axis_text_x = p9.element_text(angle=90))
)
```



Out[33]: <ggplot: (128972234360)>

In [34]:
```
my_custom_theme = p9.theme(axis_text_x = p9.element_text(color="grey", size=10,
                                                         angle=90, hjust=.5),
                           axis_text_y = p9.element_text(color="grey", size=10))
(p9.ggplot(data=surveys_complete,
           mapping=p9.aes(x='factor(year)'))
    + p9.geom_bar()
    + my_custom_theme
)
```



Out[34]: <ggplot: (-9223371907883002820)>

In [36]:
```
my_plot = (p9.ggplot(data=surveys_complete,
           mapping=p9.aes(x='weight', y='hindfoot_length'))
    + p9.geom_point()
)
```

In [37]:
```
my_plot.save("scatterplot.png", width=10, height=10, dpi=300)
```

```
E:\Install\anaconda\lib\site-packages\plotnine\ggplot.py:721: PlotnineWarning: Savin
g 10 x 10 in image.
E:\Install\anaconda\lib\site-packages\plotnine\ggplot.py:722: PlotnineWarning: Filen
ame: scatterplot.png
```

In [ ]:
```

```