Analysis: Himalayan Bird Abundance Data

ENTER

OVERVIEW OF DATA TABLES

01

02

03

04

Himalayan Bird Abundances

- 38 specific locations (“X1”)
- 300+ different bird species
- How many birds from each species observed at each location

Himalayan Forest Metadata

- 38 specific locations (“grid”) → geographic and bioclimatic data
- Elevation, latitude/longitude coordinates, region (Eastern or Western)
- Maximum temperatures, annual temperature range, annual precipitation, and precipitation seasonality

01

02

03

04

STEP-BY-STEP SOLUTIONS

Detailed solutions, along with it's codes!

QUESTION 1&2

Shared information between each table

- Columns “X1” in birds dataset and “grid” in the forest dataset = specific location
- Linked by how forest_metadata is a detailed df that expands on the sites in the birds df

Overcoming Challenges with Data

- Renamed “X1” column to “grid” to use the inner_join function and merge tables
- Birds data set: “wide” dataframe → “long” dataframe
- Used filter(n>0) to remove all data points where zero birds were found at a location

Good Tidy Data Practices

- Creating a new column for units or a README file to explain variables
- Creating a table is necessary

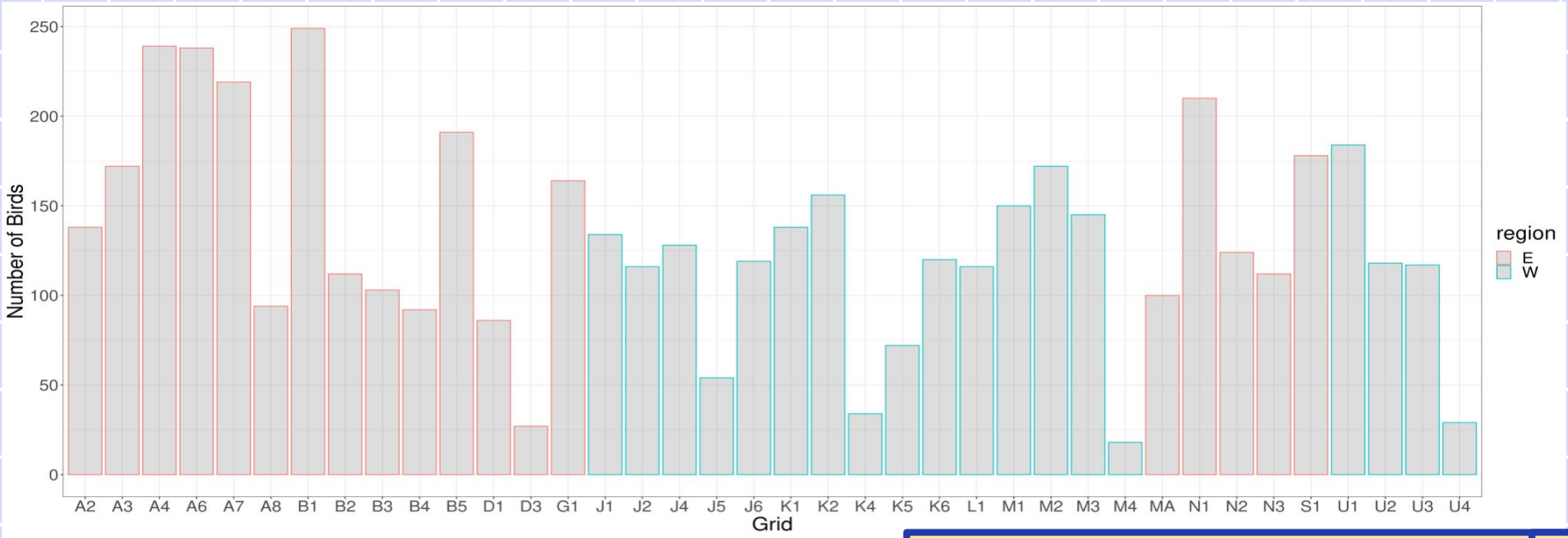
How many individual birds were sighted in each grid? How many individual birds were sighted overall in each region?

Number of Individual Birds Found in Each Region

Region	Total Birds
East	2848
West	2120

```{r}
birds_in_grid <- data_joined_1 %>%
  group_by(grid) %>%
  summarize(n_birds=sum(n))

birds_in_region <- data_joined_1 %>%
  group_by(region) %>%
  summarize(n_birds=sum(n))
```



# of birds observed in each grid and region

```
ind_plot <- data_merge %>%
  group_by(grid, region) %>%
  summarise(birds_individual=sum(n)) %>%
  ggplot(mapping=aes(x=grid,y=birds_individual,color=region))+
  geom_col(alpha = .2) +
  labs(y = "Number of Birds", x = "Grid") +
  theme_bw() +
  theme(text = element_text(size = 25))

ggsave("ind_plot.png",ind_plot , width = 25, height = 10)
```

How many bird species were present in each grid? How many bird species were present overall in each region?

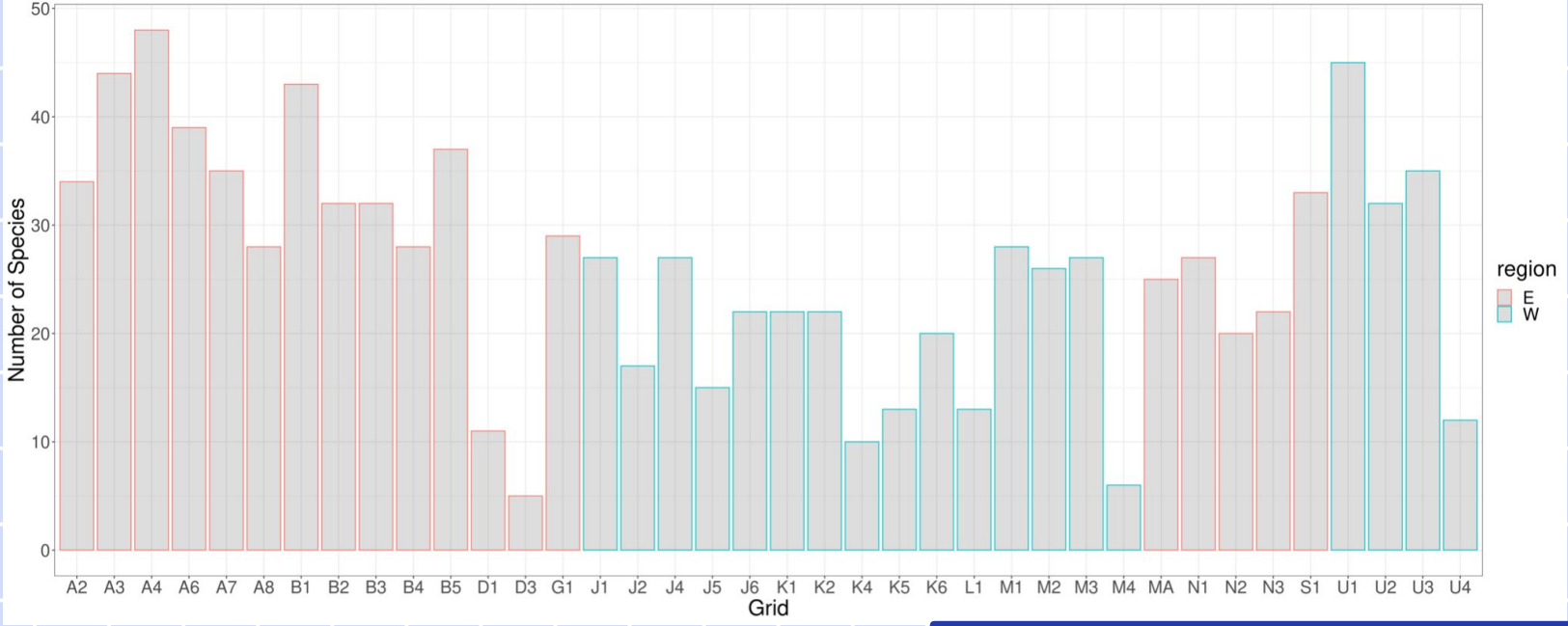
grid	n_species
A2	34
A3	44
A4	48
A6	39
A7	35
A8	28
B1	43
B2	32
B3	32
B4	28
B5	37
D1	11
D3	5
G1	29

Total number of bird species found in each grid and region

Region	Total Species
East	216
West	163

```r
species_in_grid <- data_joined_1 %>%
  group_by(grid) %>%
  summarize(n_species=n_distinct(species))

species_in_region <- data_joined_1 %>%
  group_by(region) %>%
  summarize(n_species=n_distinct(species))
```



# of species observed
in each grid and region

```
ind_plot_2 <- data_merge %>%
  group_by(grid, region) %>%
  summarise(species=n_distinct(species)) %>%
  ggplot(mapping=aes(x=grid,y=species ,color=region))+
  geom_col(alpha = .2) +
  labs(y = "Number of Species", x = "Grid") +
  theme_bw() +
  theme(text = element_text(size = 25))

ggsave("ind_plot_2.png",ind_plot_2 , width = 25, height = 10)
```

Which grid had the highest number of bird species in each region? What is the elevation and precipitation of each of those grid locations?

Grid	Elevation	Annual Precipitation	Total species
A4 (East)	1243.3	2209	48
U1 (West)	435.0	*NA*	45

3 Most Common Species in the **East**

- *Leiothrix argentauris* (n=30)
- *Schoeniparus cinerea* (n=30)
- *Phylloscopus xanthoschistos* (n=18)

3 Most Common Species in the **West**

- *Cyornis rubeculoides* (n=20)
- *Leiothrix lutea* (n=12)
- *Copsychus saularis* (n=10)
- *Phylloscopus xanthoschistos* (n=10)

```r
data_joined_1 %>%
  group_by(grid,region) %>%
  summarize(n_species=n_distinct(species)) %>%
  arrange(desc(n_species)) %>%
  head(n=2)

data_joined_1 %>%
  filter(grid=="A4"|grid=="U1") %>%
  group_by(grid,elevation,annual_precip) %>%
  summarize(n_species=n_distinct(species))

max_species_by_n <- data_joined_1 %>%
  filter(grid=="A4"|grid=="U1") %>%
  select(grid,region,species,n) %>%
  group_by(grid) %>%
  slice_max(n=3,order_by=n)
```

Within each region, which grid location had the least amount of bird species observed? Which species were observed in each of those grid locations?

Grid	Region	Elevation	Total species
D3	East	4094.3	5
M4	West	3730.0	6

```{r}
data_joined %>%
  group_by(grid, region) %>%
  summarize(n_species=n_distinct(species)) %>%
  arrange(n_species) %>%
  head(n=2)

data_joined %>%
  filter(grid=="D3"|grid=="M4") %>%
  select(grid, region, species, elevation, n) %>%
  group_by(region)
```

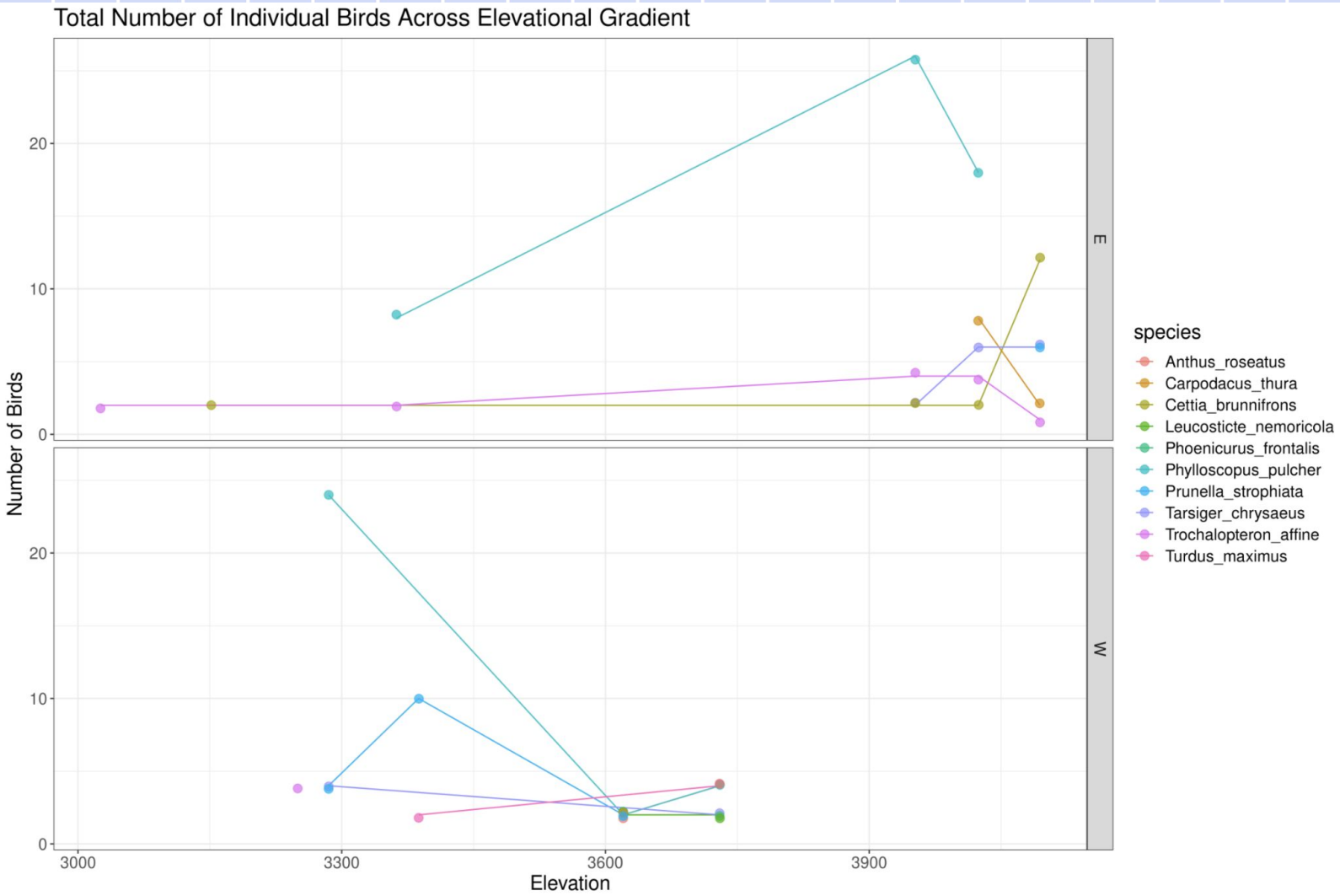
Species Observed in **D3**

- *Tarsiger chrysaeus* (n=6)
- *Cettia brunnifrons* (n=12)
- *Trochalopteron affine* (n=1)
- *Carpodacus thura* (n=2)
- *Prunella strophiata* (n=6)

Species Observed in **M4**

- *Tarsiger chrysaeus* (n=2)
- *Phoenicurus frontalis* (n=2)
- *Turdus maximus* (n=4)
- *Phylloscopus pulcher* (n=4)
- *Leucosticte nemoricola* (n=2)
- Anthus roseatus (n=4)

```
least_species_indiv <- data_joined %>%
  filter(species=="Tarsiger_chrysaeus"|
         species=="Cettia_brunnifrons"|
         species=="Trochalopteron_affine"|
         species=="Carpodacus_thura"|
         species=="Prunella_strophiata"|
         species=="Tarsiger_chrysaeus"|
         species=="Phoenicurus_frontalis"|
         species=="Turdus_maximus"|
         species=="Phylloscopus_pulcher"|
         species=="Leucosticte_nemoricola"|
         species=="Anthus_roseatus") %>%
  ggplot(mapping=aes(x=elevation,y=n,color=species))+
  geom_line()+
  geom_point(size = 3, alpha = 0.7, position=position_jitter(h=0.255,w=0.26)) +
  facet_grid(facets=vars(region)) +
  labs(y = "Number of Birds", x = "Elevation", title = "Total Number of Individual Birds
Across Elevational Gradient") +
  theme_bw() +
  theme(text = element_text(size = 16))
```


Total Number of Individual Birds Across Elevational Gradient
Number of Birds
Elevation
E
W
0
10
20
3000
3300
3600
3900
species
Anthus_roseatus
Carpodacus_thura
Cettia_brunnifrons
Leucosticte_nemoricola
Phoenicurus_frontalis
Phylloscopus_pulcher
Prunella_strophiata
Tarsiger_chrysaeus
Trochalopteron_affine
Turdus_maximus

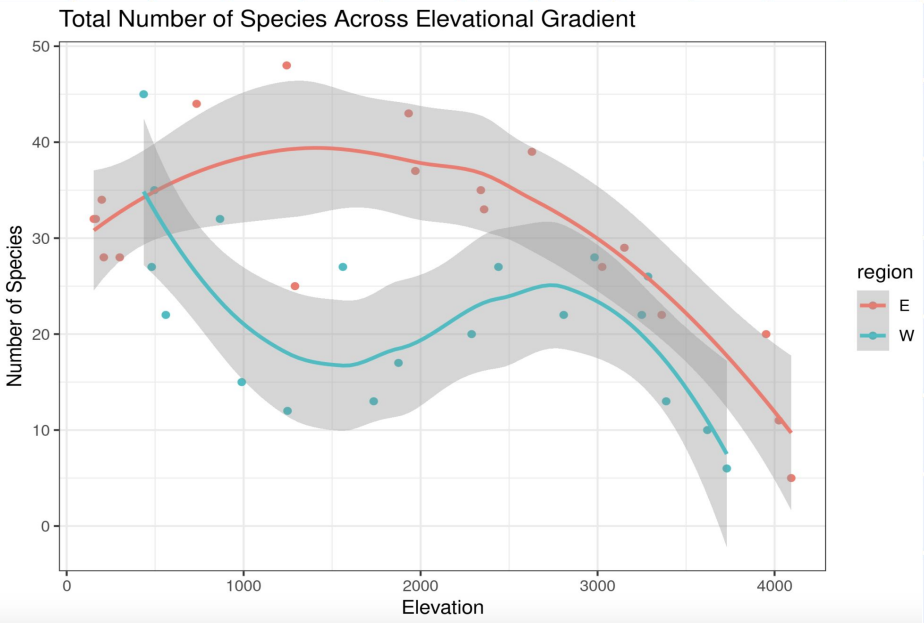
What about the total number of individual birds across the same elevational gradient?

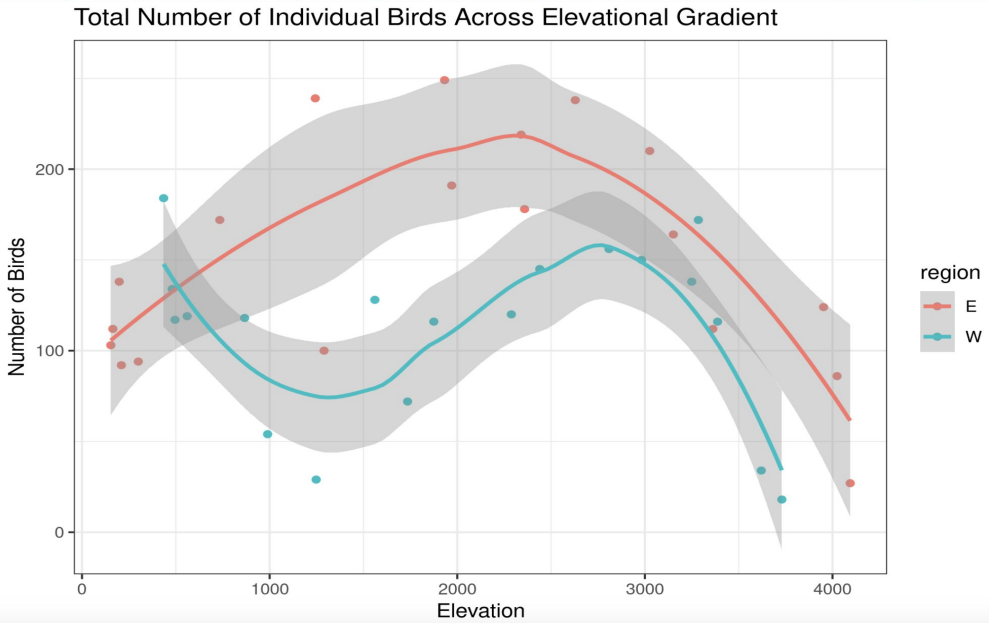
```r
#Total species across elevational gradient
data_joined %>%
  group_by(elevation) %>%
  summarise(n_species=n_distinct(species))

#Total number of individual birds across elevation gradient
data_joined %>%
  group_by(elevation)%>%
  summarise(n_individual_birds=sum(n))

#Graph of # of species in each region across elevation gradient
species_elevation <- data_joined %>%
  group_by(elevation,region) %>%
  summarise(n_species=n_distinct(species)) %>%
  ggplot(mapping=aes(x=elevation,y=n_species,color=region))+
  geom_point() +
  geom_smooth() +
  theme_bw() +
  labs(y = "Number of Species", x = "Elevation", title = "Total Number of Species Across Elevational Gradient")

#Graph of # of individual birds in each region across elevation gradient
indiv_elevation <- data_joined %>%
  group_by(elevation,region) %>%
  summarise(n_individuals=sum(n)) %>%
  ggplot(mapping=aes(x=elevation,y=n_individuals,color=region))+
  geom_point() +
  geom_smooth() +
  theme_bw() +
  labs(y = "Number of Birds", x = "Elevation", title = "Total Number of Individual Birds Across Elevational Gradient")
```


Total Number of Species Across Elevational Gradient
Number of Species
Elevation
region
E
W
0
10
20
30
40
50
0
1000
2000
3000
4000


Total Number of Individual Birds Across Elevational Gradient
Number of Birds
Elevation
region
E
W
0
100
200
0
1000
2000
3000
4000

QUESTION 4B

Using these online resources (and any others you find), please share any life history characteristics that might provide insight on the birds' status (for example, conservation status, habitat or food preference, competition, etc.)

- Over 600 species are impacted by climate change, These effects have included marked shifts in the timing of avian life cycle events, such as migration and breeding.
- ***Golden Bush Robin (Tarsiger chrysaeus)***- the global population size has not been quantified, but the species is described as rare in Pakistan, locally common in the central and eastern Himalayas, common in Bhutan and uncommon in China, stated is be in least concern.
- ***Red Breasted Parakeet (Psittacula alexandri)***- undergoing a moderately rapid population decline owing to on-going trapping pressure, persecution and habitat loss.

Bioclimatic data was provided for each site. These include max temp for warmest month, temp annual range, annual precipitation, and/or precipitation seasonality. What climate data did you find most insightful to interpreting the bird data? Please explain.

09

10

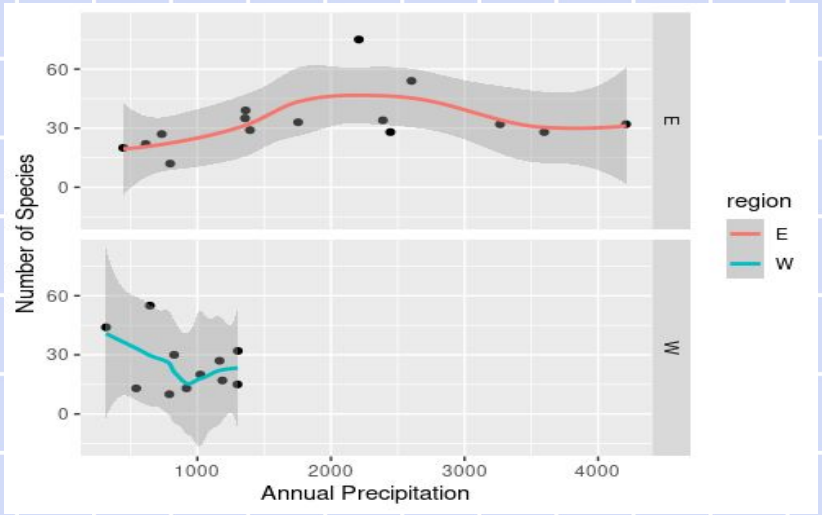
11

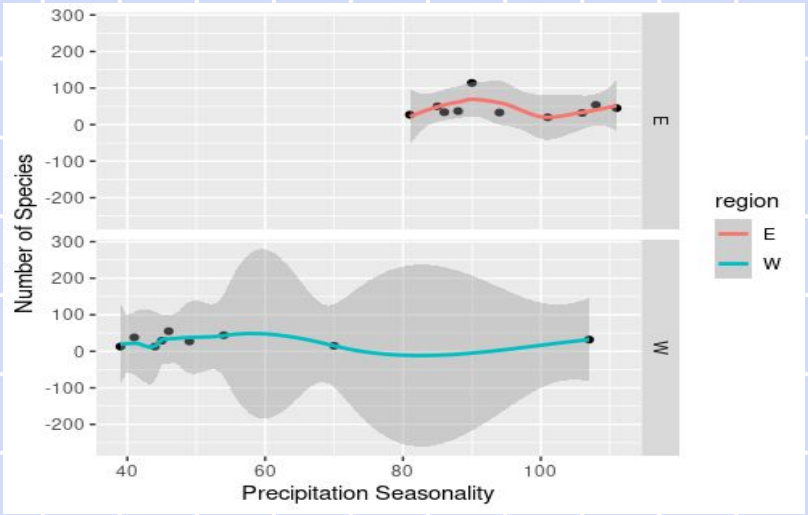
12

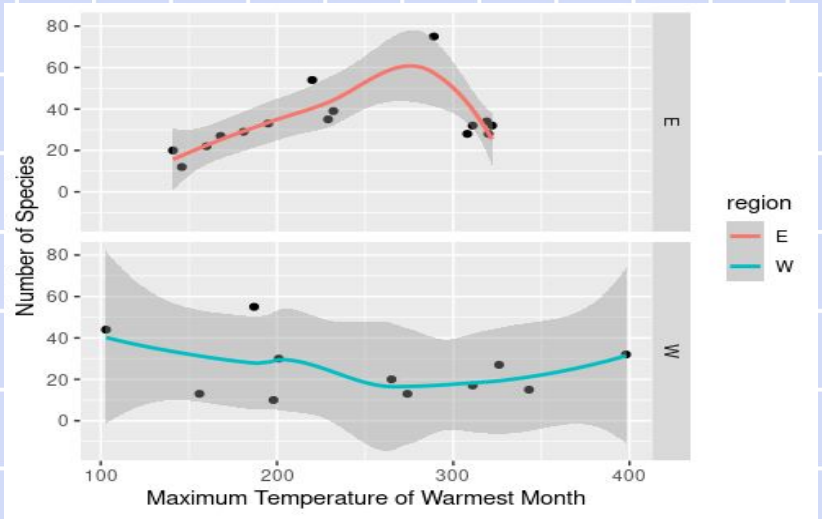
Climate Data by Region

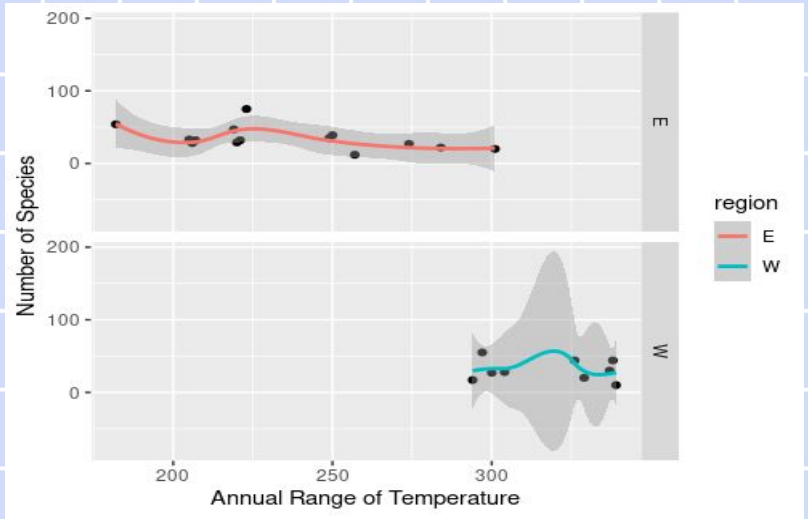
- Variables given:
  - Max Temp of Hottest Month
  - Annual Temp Range
  - Annual Precipitation
  - Precipitation Seasonality
- Looked at # of species to find trends by region
- Findings: East has more consistent temps w/ variable rainfall → biodiversity!

```
data_joined_1 %>%
  group_by(region,max_temp_of_warmest_month) %>%
  summarize(n_species = n_distinct(species)) %>%
  ggplot(mapping=aes(x=max_temp_of_warmest_month,y=n_species,color=region))+
  geom_point(color = "black") +
  geom_smooth()+
  facet_grid(facets=vars(region))+
  labs(x= "Maximum Temperature of Warmest Month", y= "Number of Species")
```


Number of Species
60
30
0
E
W
1000
2000
3000
4000
Annual Precipitation
region
E
W


Number of Species
300
200
100
0
-100
-200
E
W
40
60
80
100
Precipitation Seasonality
region
E
W


Number of Species
80
60
40
20
0
E
W
100
200
300
400
Maximum Temperature of Warmest Month
region
E
W


Number of Species
200
100
0
E
W
200
250
300
Annual Range of Temperature
region
E
W

Choose your own adventure! Present an additional visualization or trend in the data that you found useful in understanding the Himalayan region.

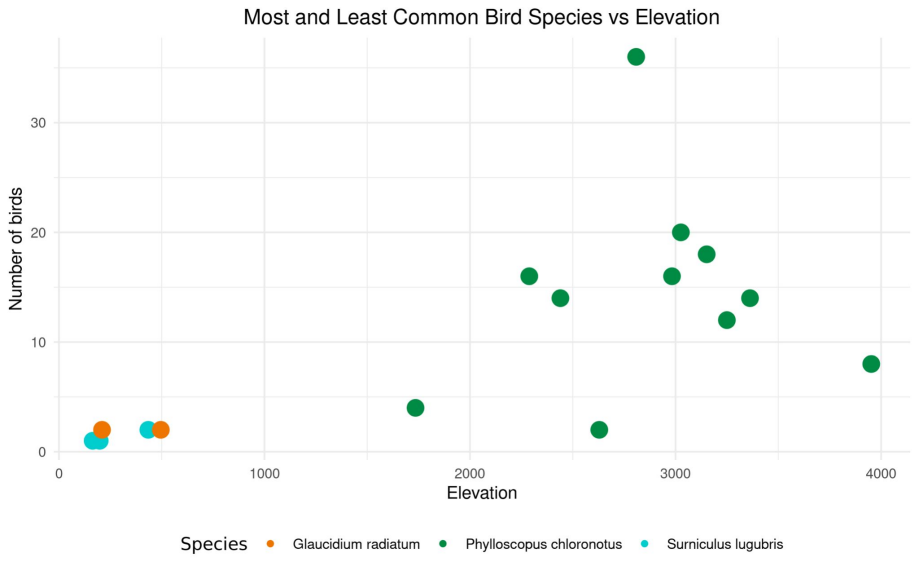
In terms of the number of birds,

- the most common species in both regions is the "**Phylloscopus chloronotus**"

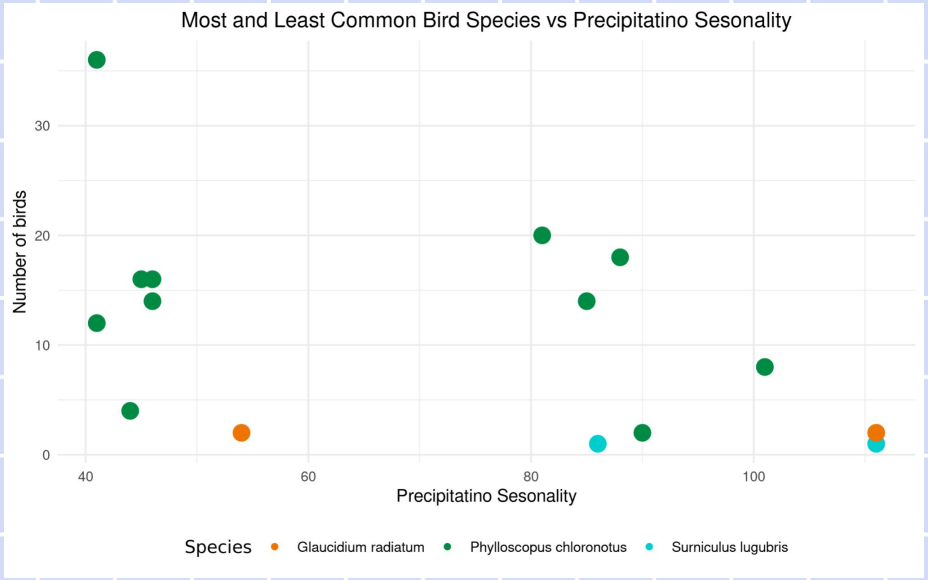
- the least common species in both regions is a tie between "**Glaucidium radiatum**" and "**Surniculus_lugubris**"

Comparing the altitude and bioclimatic conditions where these two species are found, we notice that

- elevation plays a role in the distribution –or lack thereof– of species in both regions
- precipitation seasonality doesn't seem to play as big of a role in the distribution of these species in either regions




```r
plot_1 <- data_joined_1 %>%
  filter(species=="Phylloscopus_chloronotus"|
           species=="Glaucidium_radiatum"|
           species=="Surniculus_lugubris") %>%
  ggplot(mapping=aes(x=elevation,y=n,color=species,size=12))+
  scale_colour_manual(values=c("Phylloscopus_chloronotus"="springgreen4",
                               "Glaucidium_radiatum"="darkorange2",
                               "Surniculus_lugubris"="cyan3"),
                      labels = c("Phylloscopus_chloronotus"="Phylloscopus chloronotus",
                               "Glaucidium_radiatum"="Glaucidium radiatum",
                               "Surniculus_lugubris"="Surniculus lugubris"),
                      name = "Species")+
  geom_point()+
  labs(title="Most and Least Common Bird Species vs Elevation",
       x="Elevation",y="Number of birds",size=12)+
  guides(size=F)+
  theme_minimal()+
  theme(plot.title = element_text(hjust = 0.5),
        legend.position="bottom",legend.title=element_text("Species"))
```




```r
plot_2 <- data_joined_1 %>%
  filter(species=="Phylloscopus_chloronotus"|
           species=="Glaucidium_radiatum"|
           species=="Surniculus_lugubris") %>%
  ggplot(mapping=aes(x=precip_seasonality,y=n,color=species,size=12))+
  scale_colour_manual(values=c("Phylloscopus_chloronotus"="springgreen4",
                               "Glaucidium_radiatum"="darkorange2",
                               "Surniculus_lugubris"="cyan3"),
                      labels = c("Phylloscopus_chloronotus"="Phylloscopus chloronotus",
                               "Glaucidium_radiatum"="Glaucidium radiatum",
                               "Surniculus_lugubris"="Surniculus lugubris"),
                      name = "Species")+
  geom_point()+
  labs(title="Most and Least Common Bird Species vs Precipitatino Sesonality",
       x="Precipitatino Sesonality",y="Number of birds",size=12)+
  guides(size=F)+
  theme_minimal()+
  theme(plot.title = element_text(hjust = 0.5),
        legend.position="bottom",legend.title=element_text("Species"))
```

QUESTION 7

Share what next steps you might take to investigate this data further.

- Determine which locations in each region are lacking bird species compared to other locations nearby, and help bring those numbers up by sharing the data findings and more analysis.

- Further investigate the correlation between elevation and temperature.
  - If correlated, do the two help each other or go against each other when determining the number of bird species in that location?

- Due to climate change, is temperature affecting the location to which birds travel (a new region perhaps) and/or the number of bird species who stayed in their normal location? Adding a time variable to the data could help answer this question and more.

KEY TAKEAWAYS & CLOSING STATEMENT

Regions

After researching each the eastern and western regions thoroughly, we found that there is more biodiversity among bird species in the Eastern region of the Himalayas. However when trying to research a certain species that isn't very common there is a lack of information.

THANK YOU!