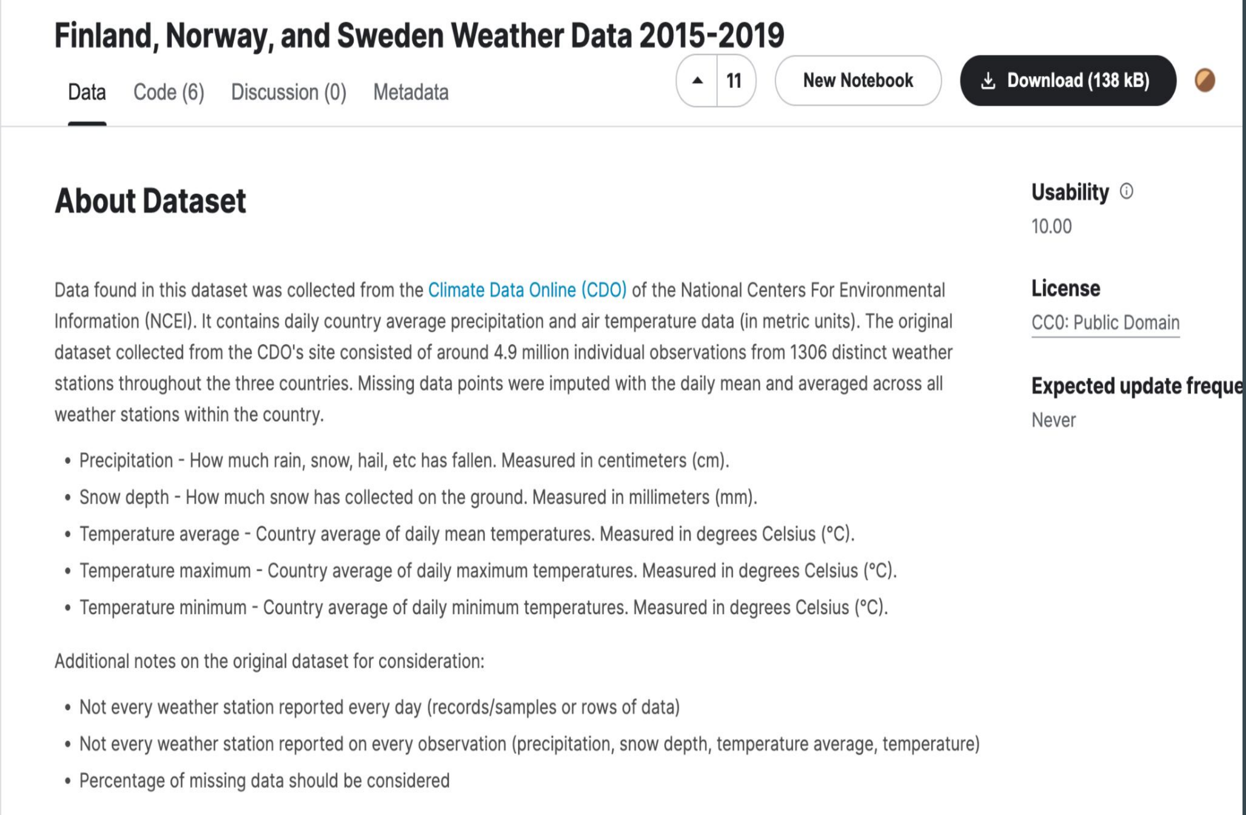
Using Simple Meteorology Data to Estimate the Wegener Bergeron Process

•••

Joel Verghese
Aos 145

Background



Finland, Norway, and Sweden Weather Data 2015-2019

About Dataset

Data found in this dataset was collected from the Climate Data Online (CDO) of the National Centers For Environmental Information (NCEI). It contains daily country average precipitation and air temperature data (in metric units). The original dataset collected from the CDO's site consisted of around 4.9 million individual observations from 1306 distinct weather stations throughout the three countries. Missing data points were imputed with the daily mean and averaged across all weather stations within the country.

- Precipitation - How much rain, snow, hail, etc has fallen. Measured in centimeters (cm).
- Snow depth - How much snow has collected on the ground. Measured in millimeters (mm).
- Temperature average - Country average of daily mean temperatures. Measured in degrees Celsius (°C).
- Temperature maximum - Country average of daily maximum temperatures. Measured in degrees Celsius (°C).
- Temperature minimum - Country average of daily minimum temperatures. Measured in degrees Celsius (°C).

Additional notes on the original dataset for consideration:

- Not every weather station reported every day (records/samples or rows of data)
- Not every weather station reported on every observation (precipitation, snow depth, temperature average, temperature)
- Percentage of missing data should be considered

Usability
10.00

License
CC0: Public Domain

Expected update freque

Never

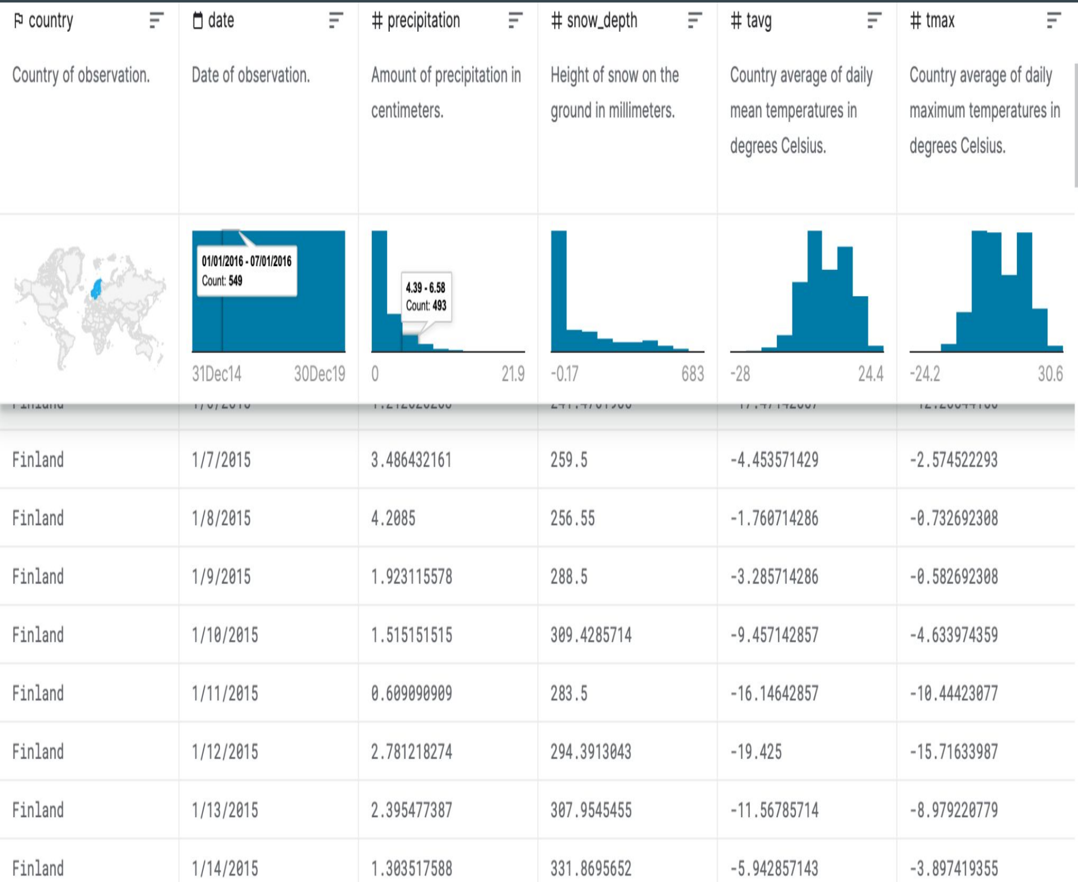
-Data can be found on Kaggle

-Google Collab was used for Code

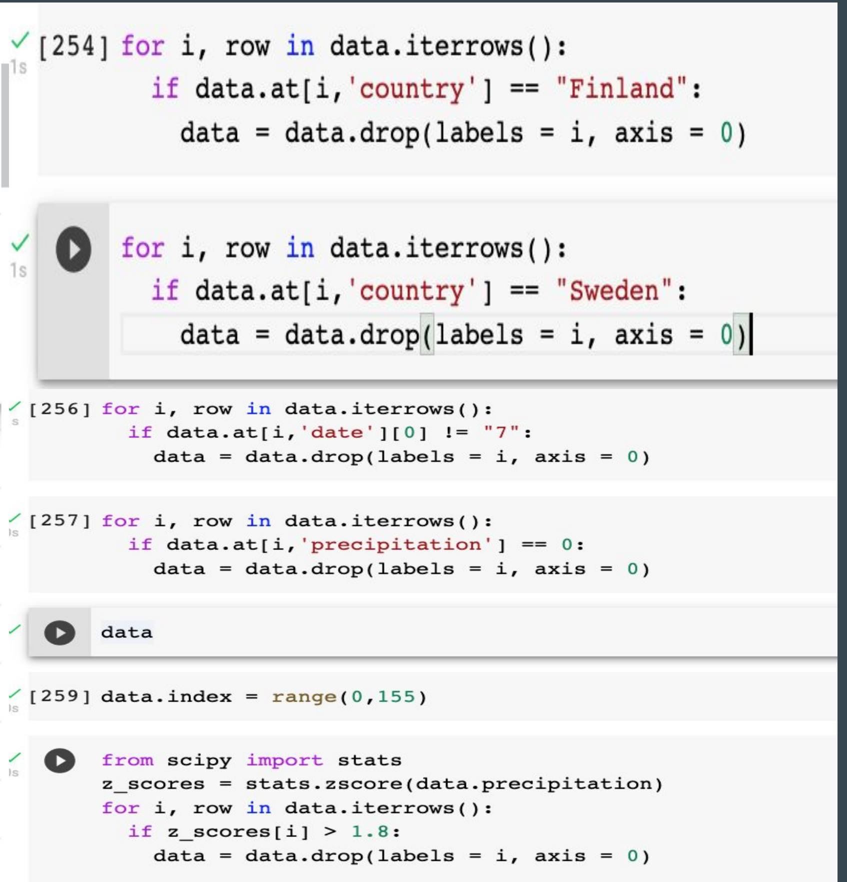
-Only looked at Norway

-Goal is to approximate when the Wegener-Bergeron Process takes effect.

Data Scrubbing



country	date	# precipitation	# snow_depth	# tavg	# tmax
Country of observation.	Date of observation.	Amount of precipitation in centimeters.	Height of snow on the ground in millimeters.	Country average of daily mean temperatures in degrees Celsius.	Country average of daily maximum temperatures in degrees Celsius.
Finland	1/7/2015	3.486432161	259.5	-4.453571429	-2.574522293
Finland	1/8/2015	4.2085	256.55	-1.760714286	-0.732692308
Finland	1/9/2015	1.923115578	288.5	-3.285714286	-0.582692308
Finland	1/10/2015	1.515151515	309.4285714	-9.457142857	-4.633974359
Finland	1/11/2015	0.609090909	283.5	-16.14642857	-10.44423077
Finland	1/12/2015	2.781218274	294.3913043	-19.425	-15.71633987
Finland	1/13/2015	2.395477387	307.9545455	-11.56785714	-8.979220779
Finland	1/14/2015	1.303517588	331.8695652	-5.942857143	-3.897419355



```
[254] for i, row in data.iterrows():
        if data.at[i,'country'] == "Finland":
          data = data.drop(labels = i, axis = 0)
```

```
for i, row in data.iterrows():
  if data.at[i,'country'] == "Sweden":
    data = data.drop(labels = i, axis = 0)
```

```
[256] for i, row in data.iterrows():
        if data.at[i,'date'][0] != "7":
          data = data.drop(labels = i, axis = 0)
```

```
[257] for i, row in data.iterrows():
        if data.at[i,'precipitation'] == 0:
          data = data.drop(labels = i, axis = 0)
```

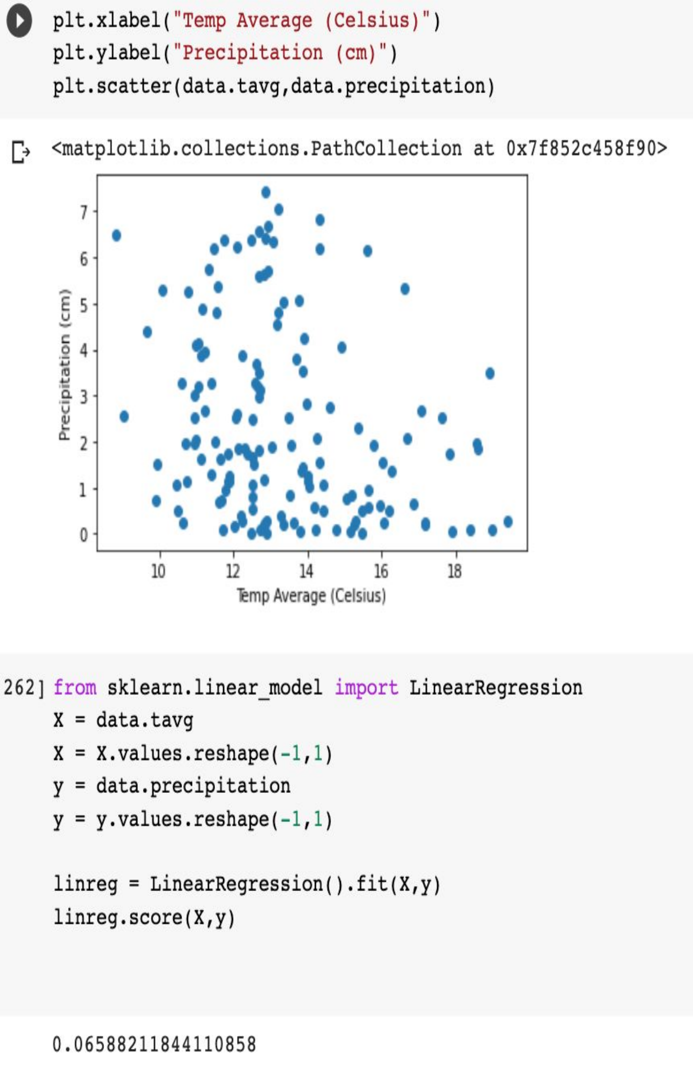
```
data
```

```
[259] data.index = range(0,155)
```

```
from scipy import stats
z_scores = stats.zscore(data.precipitation)
for i, row in data.iterrows():
  if z_scores[i] > 1.8:
    data = data.drop(labels = i, axis = 0)
```

Data Visualized

K-Means clustering



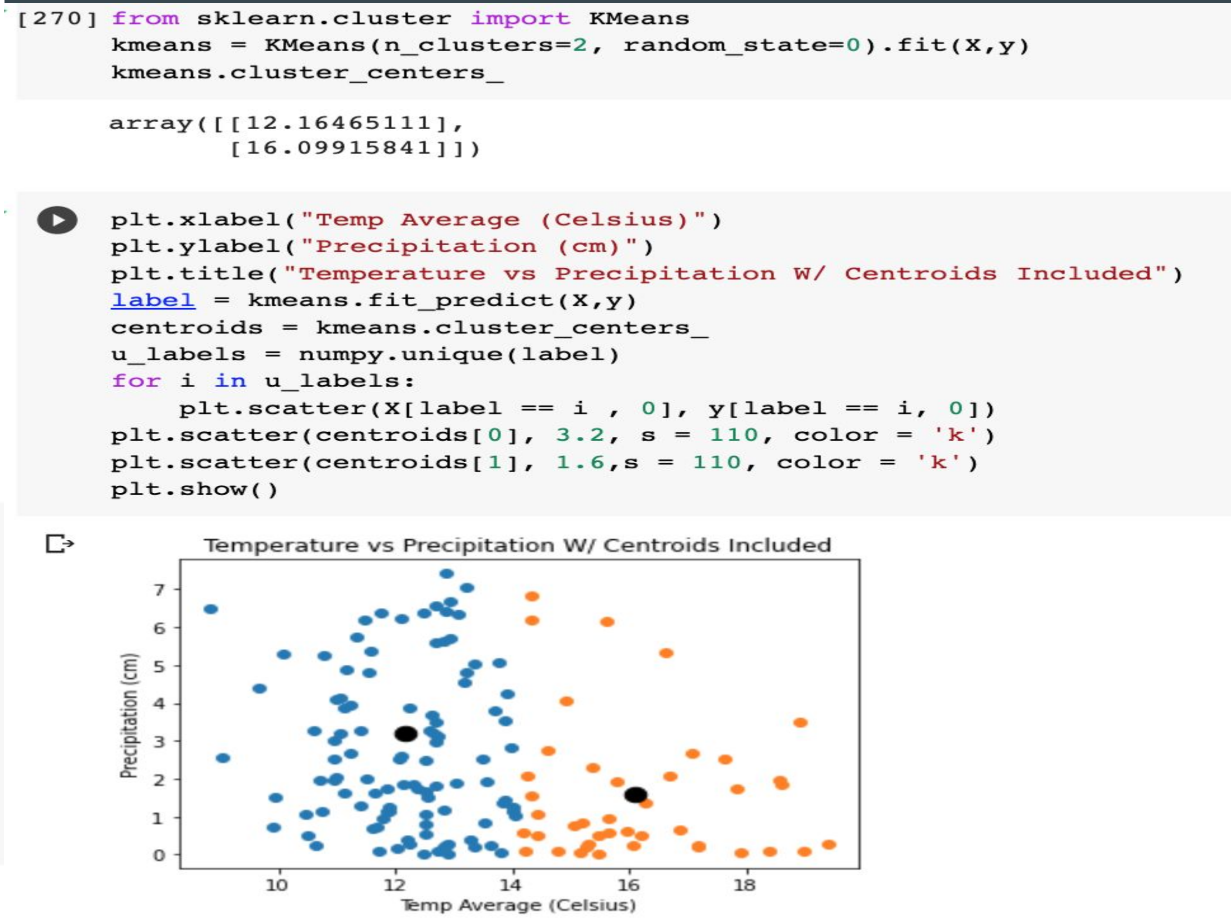
```
plt.xlabel("Temp Average (Celsius)")
plt.ylabel("Precipitation (cm)")
plt.scatter(data.tavg,data.precipitation)
```

```
<matplotlib.collections.PathCollection at 0x7f852c458f90>
```

```
262] from sklearn.linear_model import LinearRegression
     X = data.tavg
     X = X.values.reshape(-1,1)
     y = data.precipitation
     y = y.values.reshape(-1,1)

     linreg = LinearRegression().fit(X,y)
     linreg.score(X,y)
```

```
0.06588211844110858
```



```
[270] from sklearn.cluster import KMeans
      kmeans = KMeans(n_clusters=2, random_state=0).fit(X,y)
      kmeans.cluster_centers_
```

```
array([[12.16465111],
       [16.09915841]])
```

```
plt.xlabel("Temp Average (Celsius)")
plt.ylabel("Precipitation (cm)")
plt.title("Temperature vs Precipitation W/ Centroids Included")
label = kmeans.fit_predict(X,y)
centroids = kmeans.cluster_centers_
u_labels = numpy.unique(label)
for i in u_labels:
    plt.scatter(X[label == i , 0], y[label == i, 0])
plt.scatter(centroids[0], 3.2, s = 110, color = 'k')
plt.scatter(centroids[1], 1.6,s = 110, color = 'k')
plt.show()
```

Results and Explanation

.Chose to analyze the month of July because this is when temperature and furthermore air temperature most strongly correlates with precipitation.

. Location of Centroids is the end result

. Y coordinate of the Centroid was estimated

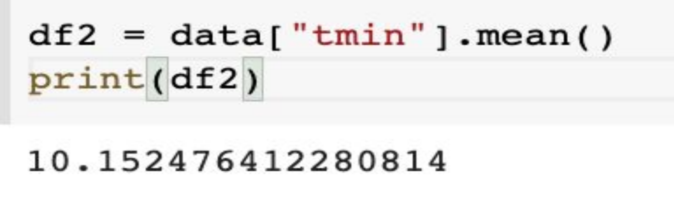
. Now this result must be compared to actual calculations

Preparation

A Possible estimation for the air temperature at which supercooled water and ice crystals form is -30 Degrees Celsius and estimate that Cirrus clouds form at 7500 meters.

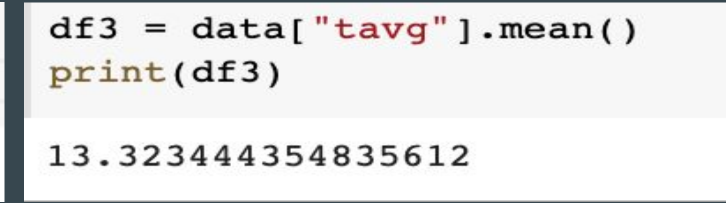
-Dewpoint Estimate:

(top left)



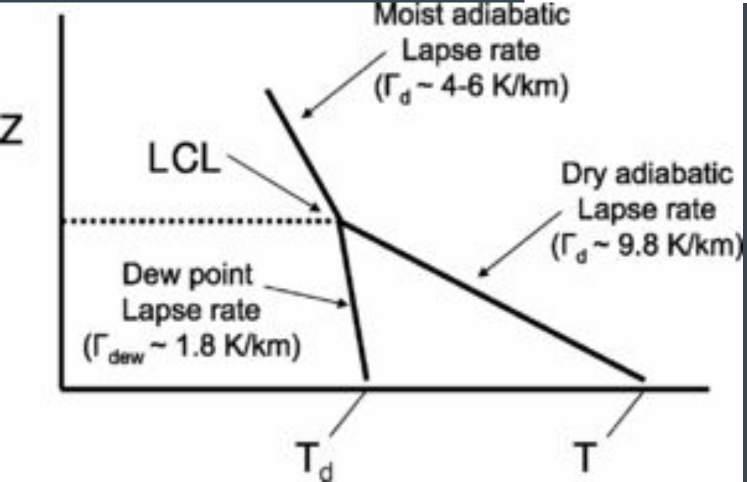
```
df2 = data["tmin"].mean()
print(df2)
```

```
10.152476412280814
```



```
df3 = data["tavg"].mean()
print(df3)
```

```
13.323444354835612
```



Lifting Condensation level diagram (Left)
An Estimate for T in diagram (Top Right)
Then by Espy's Equation: $H_{LCL} = 125\ (T\text{-}Td) = 125\ (13.3234 - 10.1525)$
= 396 meters

Estimation accurate within 1%

Moist Adiabatic lapse rate varies so estimate as:

Lapse Rates and Calculations

-Lapse rates are due to the lack of matter as altitude increases (see Ideal Gas Law)

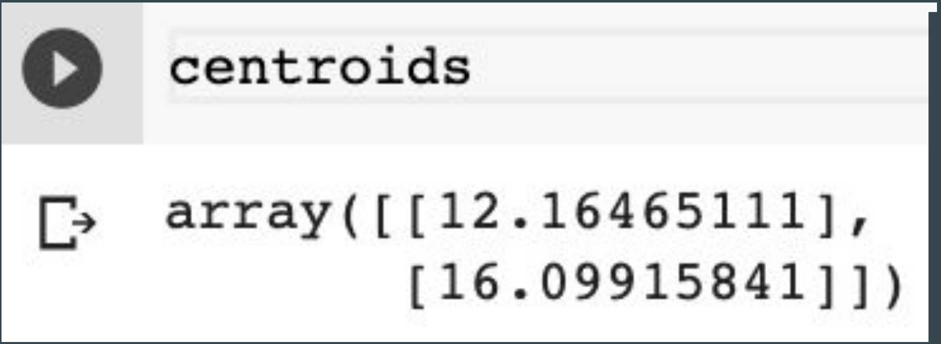
Air Temperature = T_{Ground} - 9.8 ℃/1000m (386m) – 4.0 ℃/1000m (7500m-386m)

-30 ℃ = = T_{Ground} - 9.8 ℃/1000m (386m) – 5.0 (1000m) (7500m-386m)

$\therefore T_{Ground} = 9.3258$ ℃

Comparing Results

-Data Analysis vs. Simple Physical Calculation



```
centroids
```

```
array([[12.16465111],
       [16.09915841]])
```

Air Temperature = T_{Ground} - 9.8 °C/1000m (386m) – 4.0 °C/1000m (7500m-386m)

-30 °C = = T_{Ground} - 9.8 °C/1000m (386m) – 5.0 (°C/1000m) (7500m-386m)

$\therefore T_{Ground} = 9.3258$ °C

-Both simple meteorology data and physical calculation do not account for wind speed
-Atmospheric temperature erratic
-Not easy to define the significance of a centroid
-Dew point estimate inaccurate

-Difference = 12.1647-9.3258 = 2.8389 ℃

Sources and Code

https://colab.research.google.com/drive/1l8G5LM4HOvG6U5s6xoG6Jo5pyV0nYL9J?usp=sharing

- Python on Google Collab

Works Cited

"Moist Adiabatic Lapse Rate." *NWCG*, https://www.nwcg.gov/term/glossary/moist-adiabatic-lapse-rate#:~:text=Rate%20varies%20according%20to%20the,degrees%20C%20per%201000%20meters).

Romps, David M. "Exact Expression for the Lifting Condensation Level." *AMETSOC*, American Meteorological Society, 1 Dec. 2017, https://journals.ametsoc.org/view/journals/atsc/74/12/jas-d-17-0102.1.xml.

Wurdits, Adam. "Finland, Norway, and Sweden Weather Data 2015-2019." *Kaggle*, 11 Jan. 2022, https://www.kaggle.com/datasets/adamwurdits/finland-norway-and-sweden-weather-data-20152019?phase=passwordReset&token=p6cVfL&returnUrl=%2Fdatasets%2Fadamwurdits%2Ffinland-norway-and-sweden-weather-data-20152019%3Fresource%3Ddownload&userId=9435347&userHasGoogleLogin=false&email=forjoelv%40g.ucla.edu.