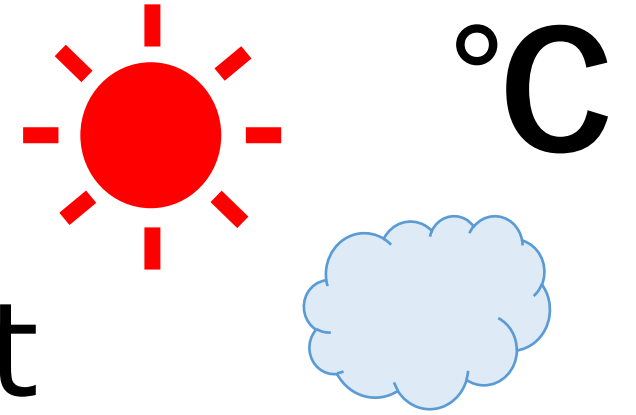


“Ski” forecast based on whether forecast

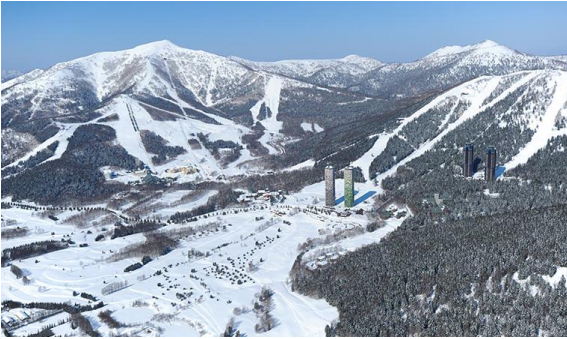


Harumi OZAKI
School of science Nagoya University

Problem Statement

There are a lot of good ski resorts in Japan.

Hokkaido Tomamu



<https://www.snowtomamu.jp/winter/ski/>

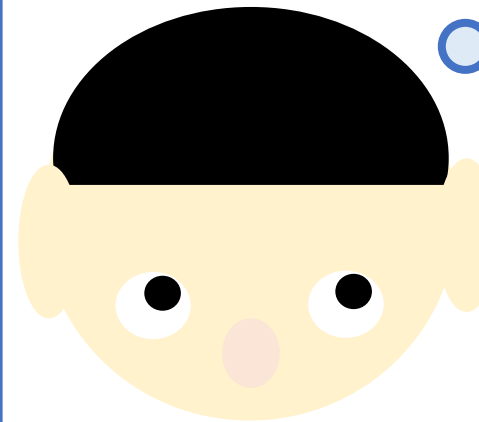


[妙高高原・赤倉温泉スキー場【公式】 \(akakura-ski.com\)](#)

Myoko Plateau, Akakura Onsen Ski Resort

However,

- Ski resorts are not always located close to our house.
- Japanese ski resorts are also popular among foreign tourists.



1 January

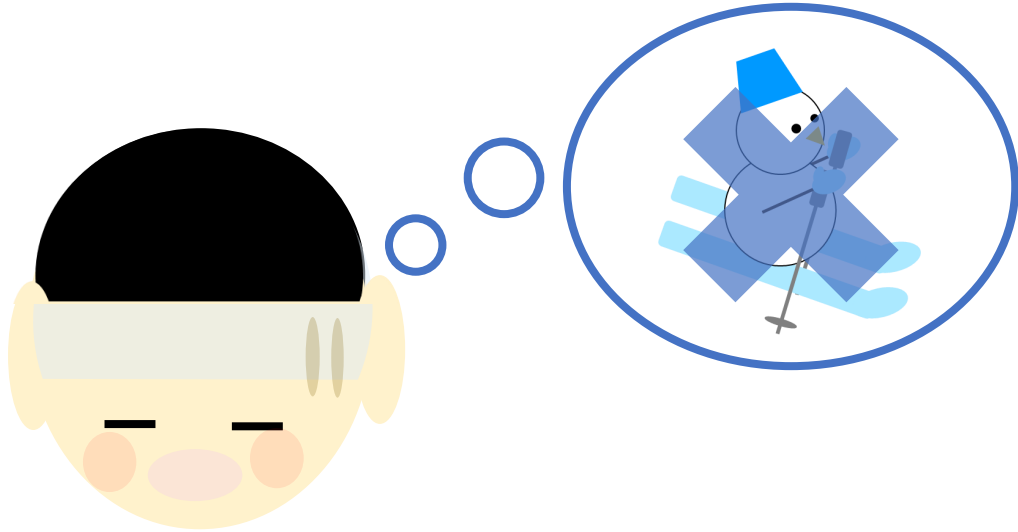
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
19	20	21	22	23	24	25

?

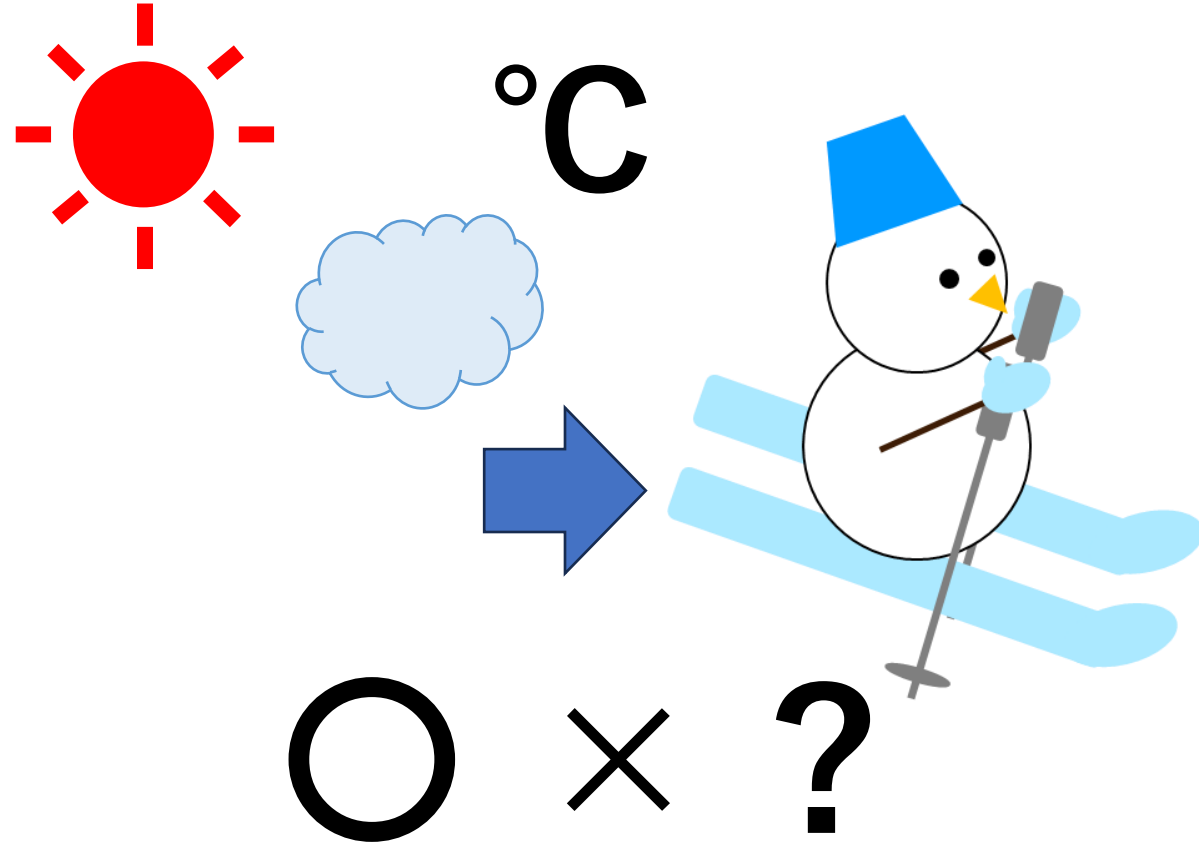
So we have to make a plan when we go to ski.

Problem Statement

When the warm weather continues, it is no longer possible to ski even if we make a reservation.



We want to prevent such a situation.



So I propose to make **"ski" forecasts based on weather forecasts.**

Data Collection

The data items I used are as follows.
I used data from **1985 to 2023**.

	年月日	最高気温(°C)	最低気温(°C)	平均気温(°C)	降水量の合計(mm)	日照時間(時間)	最深積雪(cm)				
0	1985/1/1	0.2	-3.3	-1.9	5.0	3.5	122.0				
1	1985/1/2	-0.4	-11.9	-4.1	17.0	2.1	138.0				
2	1985/1/3	0.1	-10.6	-5.4	0.0	5.5	126.0				
3	1985/1/4	-0.8	-4.0	-2.2	25.0	3.2	151.0				
4	1985/1/5	-2.7	-4.5	-3.5	53.0	1.2	205.0				
5	1985/1/6	-3.0	-5.2	-3.8	17.0	3.2	198.0				
6	1	年月日	平均気温(°C)	平均気温	平年値(°C)	降水量の合計(mm)	降水量の合計	平年値(mm)	最深積雪(cm)	最深積雪	平年値(cm)
7	1 0	1985/1/1	-1.9		-0.9	5.0		6.2	122.0		46
8	1 1	1985/1/2	-4.1		-1.0	17.0		6.2	138.0		48
9	1 2	1985/1/3	-5.4		-1.1	0.0		6.2	126.0		50
	3	1985/1/4	-2.2		-1.1	25.0		6.1	151.0		52
	4	1985/1/5	-3.5		-1.2	53.0		6.1	205.0		54
	5	1985/1/6	-3.8		-1.2	17.0		6.1	198.0		56
	6	1985/1/7	-4.3		-1.3	36.0		6.0	231.0		58
	7	1985/1/8	-9.6		-1.4	1.0		5.9	208.0		60
	8	1985/1/9	-10.6		-1.4	0.0		5.9	180.0		62
	9	1985/1/10	-2.0		-1.5	0.0		5.8	167.0		63

From the JMA website,
I downloaded some meteorological data
obtained from a station
in Tateyama, Nagano Prefecture.

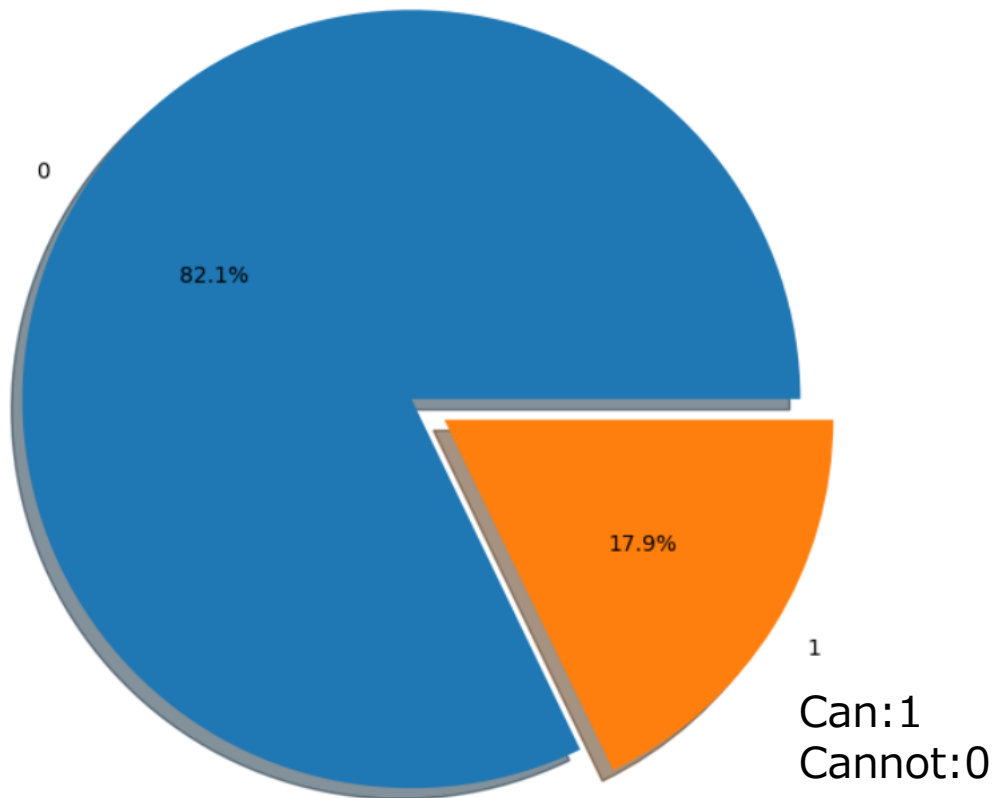
変数(variable)	定義(Definition)
年月日	date
最高気温(°C)	maximum temperature
最低気温(°C)	minimum temperature
平均気温(°C)	average temperature
降水量の合計(mm)	Precipitation totals
日照時間(時間)	daylight hours
最深積雪(cm)	Deepest Snow

変数(variable)	定義(Definition)	備考(remarks)
年月日	date	平年値(Normal value)
平均気温(°C)	average temperature	平年値(Normal value)
降水量の合計(mm)	Precipitation totals	平年値(Normal value)
最深積雪(cm)	Deepest Snow	平年値(Normal value)

Data Exploration and Visualization

I defined that “ [Deepest snow] $\geq 50\text{cm}$ ” = be able to ski

Can we ski or not?



Data Exploration and Visualization

It was difficult to download the all items so I get a part of them.

Specifically, I excluded items with many missing values or clearly meaningless items for ski in advance from the download items.

変数(variable)	定義(Definition)
年月日	date
最高気温(°C)	maximum temperature
最低気温(°C)	minimum temperature
平均気温(°C)	average temperature
降水量の合計(mm)	Precipitation totals
日照時間(時間)	daylight hours
最深積雪(cm)	Deepest Snow

変数(variable)	定義(Definition)	備考(remarks)
年月日	date	平年値(Normal value)
平均気温(°C)	average temperature	平年値(Normal value)
降水量の合計(mm)	Precipitation totals	平年値(Normal value)
最深積雪(cm)	Deepest Snow	平年値(Normal value)

Data Exploration and Visualization

The missing values of each items are as follows.

年月日	0	年月日	0
最高気温(°C)	15	平均気温(°C)	15
最低気温(°C)	15	平均気温 平年値(°C)	0
平均気温(°C)	15	降水量の合計(mm)	7
降水量の合計(mm)	7	降水量の合計 平年値(mm)	0
日照時間(時間)	11	最深積雪(cm)	4800
最深積雪(cm)	4800	最深積雪 平年値(cm)	0

- "deepest snow" : **filled in with 0**

To see the data, it is not measured during seasons when there is no snowfall .

- "maximum temperature", "minimum temperature", "average temperature",
"Precipitation total", "daylight hours" : **filled in with one previous value**

The rest of the data is time series data.

Model Building

I did a Logistic Regression Analysis.

-What is **Logistic Regression Analysis** about?

$$S(x) = \frac{1}{1 + e^{-x}} \quad \rightarrow \quad S(\theta^T x) = \frac{1}{1 + e^{-\theta^T x}}$$

Sigmoid Function

Varying the value of θ^T changes the shape of the sigmoid function.
Find the θ^T at which the graph fits the data.

How to get best θ^T ?

Maximum-Likelihood Estimation

$$l(\theta) = \log L(\theta) = \sum_{j=1}^n y^{(j)} \log h_{\theta}(x^{(j)}) + (1 - y^{(j)}) \log(1 - h_{\theta}(x^{(j)}))$$

2D

Evaluation

I used package to make a model from `sklearn.linear_model`.

It's because I couldn't found the way to make 3D or more model mathematically.

① I use 5 items to learn and test.

Training data & Test data

最高気温(°C)	maximum temperature
最低気温(°C)	minimum temperature
平均気温(°C)	average temperature
降水量の合計(mm)	Precipitation totals
日照時間(時間)	daylight hours

accuracy (train) : 0.899
accuracy (test) : 0.898

② I use 2 Normal value data to predict.

Training data

最高気温(°C)	maximum temperature	
最低気温(°C)	minimum temperature	
平均気温(°C)	average temperature	平年値(Normal value)
降水量の合計(mm)	Precipitation totals	平年値(Normal value)
日照時間(時間)	daylight hours	

Test data

accuracy (train) : 0.899
accuracy (test) : 0.898

However, It's not good to predict because I use the data on the day's weather.

Evaluation and Conclusion

Odds

1 (be able to ski) is e^x times more likely to occur than 0 (not be able to).

-The higher the odds value, the greater the role of that variable.

変数(variable)	定義(Definition)	オッズ(Odds)
最高気温(°C)	maximum temperature	0.78322372
最低気温(°C)	minimum temperature	0.75249789
平均気温(°C)	average temperature	1.13408891
降水量の合計(mm)	Precipitation totals	1.03469297
日照時間(時間)	daylight hours	1.05449789

This model successfully predict about 90% whether we can ski or not based on weather forecast.

