Lionel Perret
Route de Diesse 8
2518 Nods
Lionel.Perret.76@gmail.com
+41 76 2311093

Data Science Project

Prediction of Avalanche Danger Levels with Meteorological Data

Conceptual Design Report

16 October 2020

Abstract

This data science project analyses the data set of 13'918 avalanches, recorded in the Davos area from 1998 to 2019 and combine it with the meteorological data of the Weissfluhjoch station, above Davos. The avalanche data set includes information on the size, the type, and the number of avalanches. For each day, when avalanches are recorded, the official avalanche danger level is also present in this database.
From the snowfall quantity, the wind and the temperature data from the weather dataset, a binary variable describing the *Critical Fresh Snow* limit is created, to investigate if more avalanches were recorded with those conditions. The results showed that this variable gives an indication that the number of avalanches per day increased but also that other LOCAL meteorological parameters play an important role as well.

The final objective of this project (module 3) is to predict the avalanche danger levels, with the meteorological parameter available in the Weissfluhjoch data set and to investigate the degree of accuracy of this prediction.

Table of Contents

Abstract 0
Table of Contents 1
1. Project Objectives 2
  Project - Module 1 - Objectives 2
  Project - Module 2 – Objectives 3
  Project – Module 3 – Objectives 3
2. Methods 4
  Infrastructure 4
  Tools 4
  Python Libraries 4
3. Data 5
  Data Set 1 Snow avalanche data Davos, Switzerland, 1999-2019 5
  Data Set 2 Weissfluhjoch Meteorological and snowpack measurement 1999-2017 7
4. Metadata 9
5. Data Quality 10
  Meteorological measurements accuracy 10
  Geographic position differences in avalanches and meteorological measurements 10
  Manually recorded Variables in avalanche dataset 10
6. Data Flow 11
7. Data Model 12
  Conceptual 12
  Logical 12
  Physical 12
8. Risks 13
  Recording mistakes risk 13
  Analysis mistakes risk 13
  Data collection risk (Project Module 3) 13
9 Preliminary Studies – Module 2 14
10. Conclusion 18
  Project Module 1 18
  Project Module 2 18
  Outlook Project Module 3 18
  Broader outlook towards “LOCAL” Avalanche Danger Level Prediction 19
11. Acknowledgements 20
12. References 20
13. Table list 20
14. Figure list 20

1. Project Objectives

Project - Module 1 - Objectives

The main activity for the project of this first module will be to get familiar with the 2 datasets used and then to achieve the following goals:

- Pre-process the avalanche dataset and start to describe statistically the avalanches that were recorded: Number of avalanches, types, sizes, causes, …

- Pre-process the meteorological dataset and create a binary variable *Critical Fresh Snow* condition for each day, using the definition of the Institute for Snow and Avalanche Research, Davos (SLF) about "Critical amount of new snow reached" **[1]**:

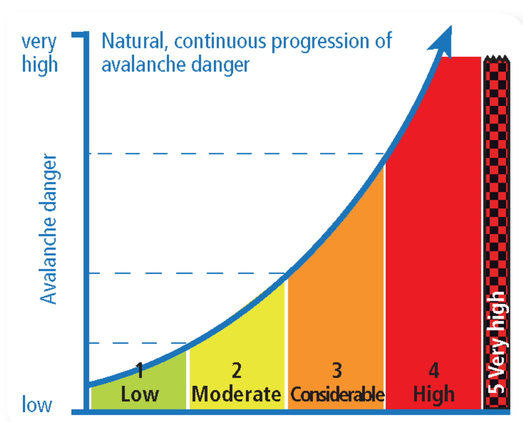
FIGURE 1.: CRITERIA THAT DEFINES CRITICAL FRESH SNOW CONDITIONS

***Critical Fresh Snow* = 1** ➔ **The meteorological conditions defined above are present** on that day
***Critical Fresh Snow* = 0** ➔**The meteorological conditions defined above are not present**

Note:
The above recommendation "Critical amount of new snow reached = at least **Considerable** avalanche danger" refers to the European avalanche danger level, more information on its definition is given in table 1. below:

TABLE 1. : AVALANCHE DANGER SCALE, EUROPEAN AVALANCHE DANGER LEVELS [2]

Level	Characteristics
5 **Very high**	**Extraordinary avalanche situation** Numerous very large and extremely large natural avalanches can be expected. These can reach roads and settlements in the valley.
4 **High**	**Very critical avalanche situation** Natural and often very large avalanches are likely. Avalanches can easily be triggered on many steep slopes. Remote triggering is typical. Whumpf sounds and shooting cracks occur frequently.
3 **Considerable**	**Critical avalanche situation** Whumpf sounds and shooting cracks are typical. Avalanches can easily be triggered, particularly on steep slopes with the aspect and elevation indicated in the avalanche bulletin. Natural avalanches and remote triggering can occur.
2 **Moderate**	**Mostly favorable avalanche situation** Warning signs can occur in isolated cases. Avalanches can be triggered in particular on very steep slopes with the aspect and elevation indicated in the avalanche bulletin. Relatively large natural avalanches are not to be expected.
1 **low**	**Generally, favorable avalanche situation** No warning signs present. Avalanches can only be triggered in isolated cases, in particular on extremely steep slopes.



Project - Module 2 – Objectives

Following the pre-processing, those two datasets will be combined and the goals for the module 2 project are:

- Analysis of number of avalanches per day with regards to avalanche danger levels.
  Here the aim is to investigate if the number of avalanches per day are increasing with the avalanche danger levels

- Investigate the influence of snow fall and other meteorological parameters on the number of avalanches per day. The binary variable *Critical Fresh Snow* created in module 1 will be used to split the data in 2 classes and the obtained distributions will be compared:

  Distribution 1: Number of avalanches per day for the day where *Critical Fresh Snow* = 1

  Distribution 2: Number of avalanches per day for the day where *Critical Fresh Snow* = 0

- Check the possibility to use the combined dataset for prediction of the avalanche danger levels with meteorological parameters

Project – Module 3 – Objectives

The final goal of this project is to use and interpret the recorded meteorological data in the Weissfluhjoch defined area to predict the avalanche danger levels within this area.
To do that the model will be trained in using the two combined datasets used in module 1-2 and the model will be used to predict the upcoming avalanche danger levels in the Davos Area of the winter 2020-2021.

2. Methods

Infrastructure

- Microsoft Surface Book 2 with Intel Core i5-8350U, 8Go, 256Go, SSD
- Windows 10 Professional, Version 10.0.19041

Tools

- Jupyter Notebook (Version 6.0.0): Interactive computing product based on the platform Anaconda 2019.07 (Python 3.7 Version [6]). Data analysis is performed in Python 3

Python Libraries

- *pandas*: Library and tool for dealing with data structures, data tables, manipulating data.
- *numpy*: Library and tool for calculating with data and arrays.
- *matplotlib*: Library and tool for the visualization of data in diagrams.
- *scipy*: Library and tool for mathematical calculations (scipy.stats: probability distributions, normality tests, regression etc.)

3. Data

Data Set 1 Snow avalanche data Davos, Switzerland, 1999-2019

Description: This dataset includes observations about all avalanches recorded in the region of Davos, Switzerland, during the winters 1998-1999 to 2018-2019 - 21 years – total 13’918 avalanches **[3]**

Variables Date, avalanche number, snow type, trigger type, maximum elevation, minimum elevation, length, width, perimeter, area, avalanche size class, avalanche activity index, avalanche danger level

Snow Type variable:

- “dry” avalanche: Type of avalanche where the snow has not melted. This type of avalanche is either with fresh powder snow or “slab” snow. “Slabs” are caused by the effect of the wind on the snow.
- “wet” avalanche: In this avalanche, the snow as melt because of the sun and/or temperature. This type of avalanche is more frequent in the second part of the winter.

Trigger Type variable

- “EXPLOSIVE”: avalanche triggered artificially, for security reasons
- “NATURAL”: avalanche triggered by natural causes
- “HUMAN”: avalanche triggered by a person unintentionally
- “UNKNOW”: Triggered cause of the avalanche is unknown

TABLE 2.: DATA FROM THE ORIGINAL AVALANCHE DATASET

	Date	Snow_type	Trigger_type	Avalanche_size_m2	Avalanche_risk_level
11394	2017-03-10	wet	EXPLOSIVE	7,022	4
11395	2017-03-10	dry	EXPLOSIVE	9,953	4
11396	2017-03-10	dry	NATURAL	3,306	4
11397	2017-03-10	dry	EXPLOSIVE	10,339	4
11398	2017-03-10	dry	HUMAN	3,925	4
11399	2017-03-10	dry	NATURAL	1,411	4

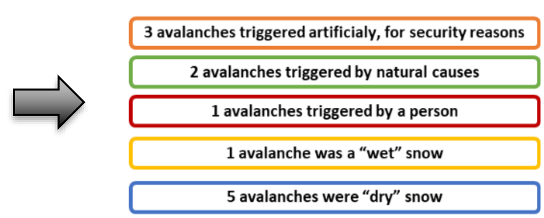


TABLE 2. gives an example of the original data. There is one line per recorded avalanche, so several avalanches recorded on the same day are represented with different rows.

Over the years, the number of Natural triggered avalanches are fluctuating and in the past 2 years, the number of "HUMAN" caused avalanches is decreasing.

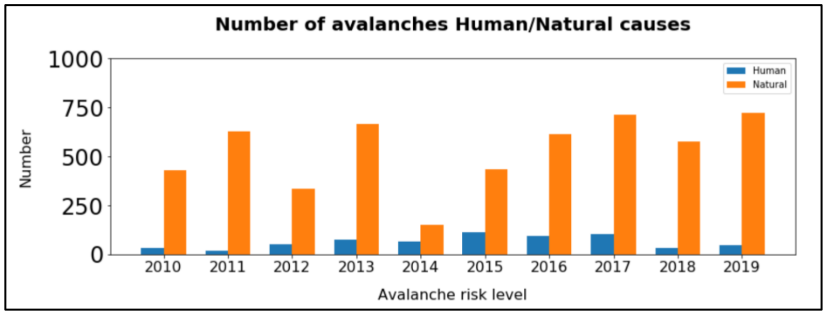


FIGURE 2.: NUMBER OF AVALANCHES WITH HUMAN/NATURAL CAUSES

In FIGURE 3. the distribution of the avalanche size has a similar shape for "wet" and "dry" avalanches and is not normal.

FIGURE 4. shows the box plot of the 2 distributions with and without outliers. The avalanche size can be very big, in 2 cases for "dry" avalanches, above 1'000'000 square meters.

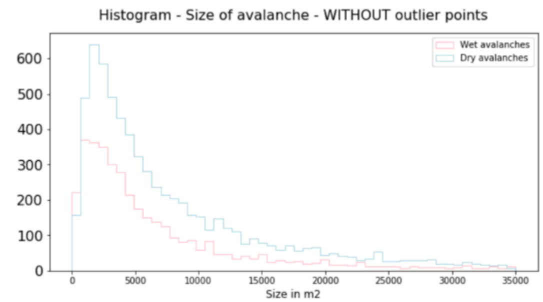


FIGURE 3.: HISTOGRAM AVALANCHE SIZE

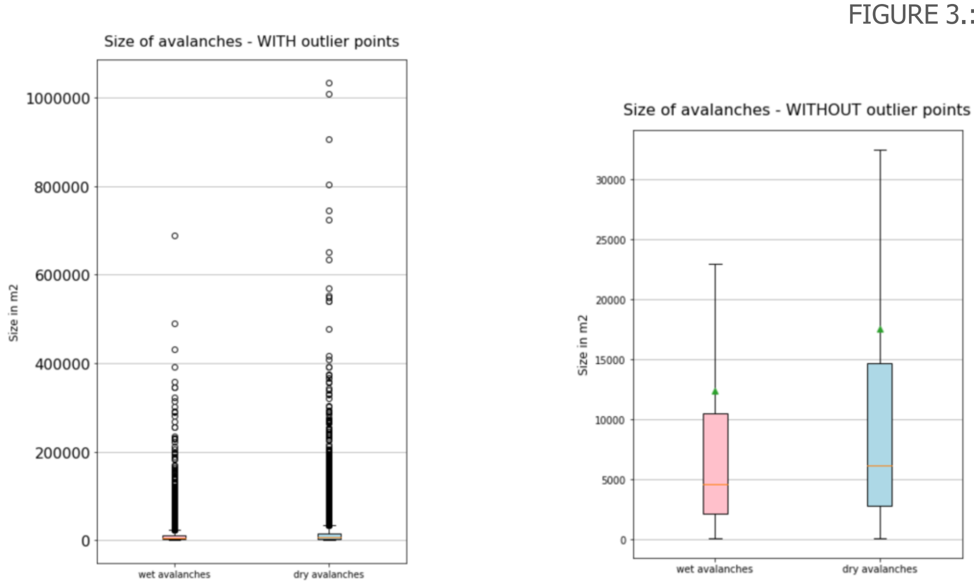


FIGURE 4.: BOX PLOT AVALANCHE SIZE

TABLE 3. below shows the statistic of "dry Avalanches.

TABLE 3.: DESCRIPTIVE STATISTIC "DRY" AVALANCHE SIZE

Statistic about "dry" avalanches size

```
Entrée [13]:  df.xs(('dry'), level = ('Snow_type')).Avalanche_size_m2.describe()

   Out[13]:   count        19.0
              mean     33,437.6
              std      75,213.0
              min       3,100.5
              25%       6,504.2
              50%      14,456.6
              75%      18,634.9
              max     338,462.5
              Name: Avalanche_size_m2, dtype: float64
```

Data Set 2 Weissfluhjoch Meteorological and snowpack measurement 1999-2017

Description: This dataset includes standard meteorological and snowpack measurements from the automatic weather station at Weissfluhjoch, Davos, Switzerland **[4]**

Variables Timestamp, air temperature, relative humidity, wind speed, wind direction, incoming short wave radiation, short wave radiation, incoming long wave radiation, outgoing long wave radiation, undercatch corrected precipitation, manually observed snow depth, once per day, automatic measured snow depth, ground temperature, surface temperature, snow lysimeter, measured snowpack runoff, snow water equivalent

TABLE 3. gives an overview of the raw data of 04.03.2017 and 05.03.2017. The meteorological parameters are measured every 30 minutes and *Snow_height* is manually measured every day between 06h and 09h:

TABLE 4.: RAW DATA FROM THE METEOROLOGICAL DATA SET

	Date	Temp	RH	Wind	ISWR	OSWR	ILWR	OLWR	Snow_height	TSS
359484	2017-03-04 07:00:00	-3.6	0.8	26.3	1.3	1.1	203.7	282.0	nan	-8.0
359485	2017-03-04 07:30:00	-3.5	0.8	25.9	27.4	22.4	204.7	281.6	nan	-8.1
359486	2017-03-04 08:00:00	-4.1	0.8	25.6	53.2	44.2	200.2	278.6	139.0	-8.8
359487	2017-03-04 08:30:00	-4.0	0.8	27.4	184.4	150.8	203.2	282.0	nan	-8.1

	Date	Temp	RH	Wind	ISWR	OSWR	ILWR	OLWR	Snow_height	TSS
359532	2017-03-05 07:00:00	-7.8	1.0	2.5	2.3	0.8	274.9	279.6	nan	-8.4
359533	2017-03-05 07:30:00	-7.6	1.0	1.4	43.9	39.4	274.3	280.1	nan	-8.3
359534	2017-03-05 08:00:00	-7.1	1.0	1.8	123.5	112.4	268.6	281.1	149.0	-8.1
359535	2017-03-05 08:30:00	-7.0	1.0	1.8	133.9	125.7	274.1	281.9	nan	-7.9

Binary variable *Critical Fresh Snow:*
The *Critical Fresh Snow* variable was created using simplified criteria defined in FIGURE 5 below.

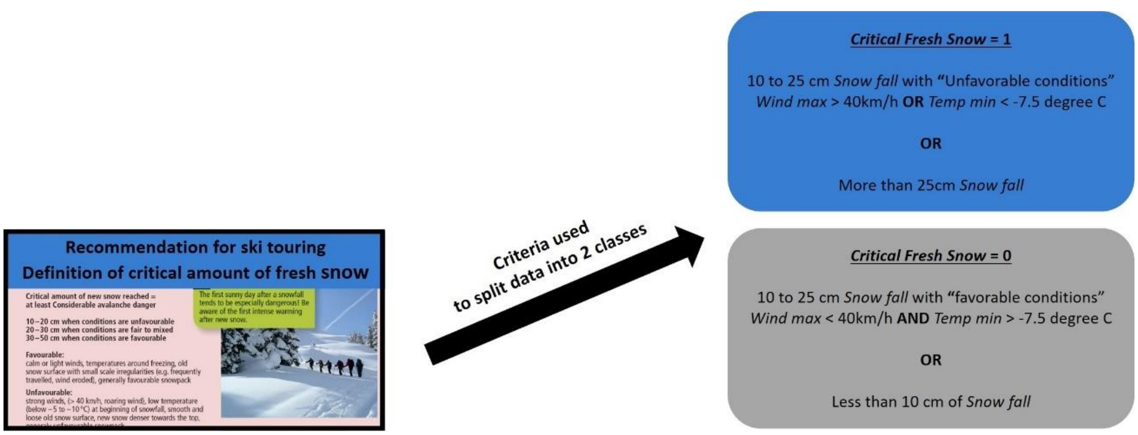


FIGURE 5.: SIMPLIFIED CRITERIA FOR THE CRITICAL FRESH SNOW VARIABLE

TABLE 4. gives an overview of the data after the pre-processing, with the *Critical Fresh Snow* variable calculated The 10cm snow falls on 04.03.2017 is obtained with the difference of the two *Snow_height* values in yellow in TABLE 3. The processed data in TABLE 4. is exported as a .CSV file in the *Meteo.ipynb* Jupyter notebook. The *Data model* chapter below gives a description of the used file names.

TABLE 5.: PROCESSED DATA FROM THE METEOROLOGICAL DATA SET (METEO.CSV FILE)

	Date	Snow_fall	Snow_fall_1	Snow10To25	SnowAbove25	Temp_min_1	TempBelow-7_5	Wind_max_3	WindAbove40	CriticalFreshSnow
6348	2017-03-02	0	0	False	False	-12.0	True	31.3	False	0
6349	2017-03-03	0	0	False	False	-7.8	True	23.0	False	0
6350	2017-03-04	10	0	False	False	-8.4	True	26.6	False	0
6351	2017-03-05	11	10	True	False	-7.1	False	29.9	False	0
6352	2017-03-06	15	11	True	False	-8.4	True	29.9	False	1
6353	2017-03-07	0	15	True	False	-10.2	True	29.9	False	1

Descriptive statistic for the daily mean temperature is presented in FIGURE 6 below.

```
Entrée [33]:  df_meteo['Temp'].describe()

   Out[33]:  count    6,480.0
             mean        -0.4
             std          7.0
             min        -23.7
             25%         -5.3
             50%         -0.2
             75%          4.9
             max         16.4
             Name: Temp, dtype: float64
```

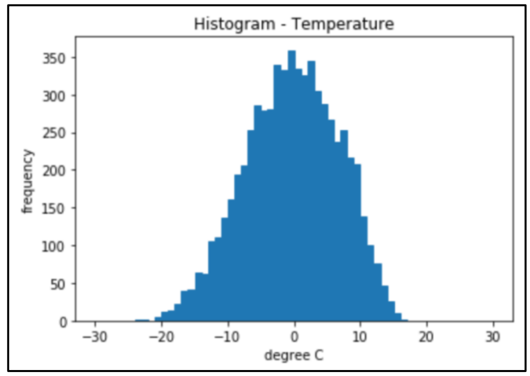


FIGURE 6.: MEAN DAILY TEMPERATURE - WEISSFLUHJOCH WEATHER STATION

Descriptive statistic for the sum of the snow fall of the last 3 days is presented in FIGURE 7 below.

```
Entrée [34]:  df_meteo['Snow_fall_3'].describe()

   Out[34]:  count    6,480.0
             mean         0.0
             std         29.5
             min       -264.0
             25%         -5.0
             50%          0.0
             75%          2.0
             max        272.0
             Name: Snow_fall_3, dtype: float64
```

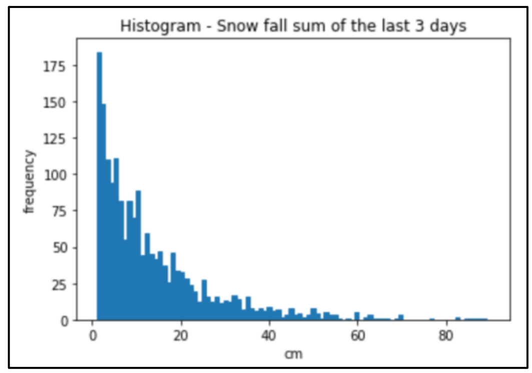


FIGURE 7.: SUM OF SNOW FALL OF LAST 3 DAYS - WEISSFLUHJOCH WEATHER STATION

Descriptive statistic for the mean daily wind speed value is presented in FIGURE 8 below.

```
Entrée [48]:  df_meteo['Wind'].describe()

   Out[48]:  count    6,480.0
             mean         8.2
             std          4.0
             min          0.0
             25%          5.2
             50%          7.2
             75%         10.3
             max         32.1
             Name: Wind, dtype: float64
```

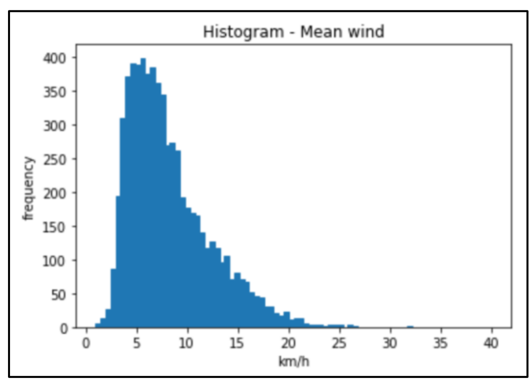


FIGURE 8.: MEAN DAILY WIND SPEED - WEISSFLUHJOCH WEATHER STATION

4. Metadata

Metadata information can be found on the following Github directory:

https://github.com/lionelperret76/M1-M2_Project/tree/main/MetaData

In this directory:

- *Meteo_metadata.txt* contains the metadata information about the meteo variables that are found in the *Meteo.csv* file

- *Avalanches_metadata.txt* contains the metadata information about the avalanches variables that are found in the *Avalanches.csv* file

5. Data Quality

Meteorological measurements accuracy

The meteorological measurements are mostly automatic. The precision of the temperature, wind, snow height measurement is not stated in the dataset information. The automatic measurement system at the Weissfluhjoch are expected to be precise enough for the analysis planned in this project.

Geographic position differences in avalanches and meteorological measurements

The meteorological measurements available in the dataset are from only one location - The Weissfluhjoch is represented by the "blue dot" in FIGURE 6. below.
All the recorded avalanches are located within the red area, around Davos. As snow fall quantity, wind speed and temperature can be quite different at the avalanches location, some kilometer away and/or at another altitude than the measured meteorological values, this could have an impact on the analysis in this project.

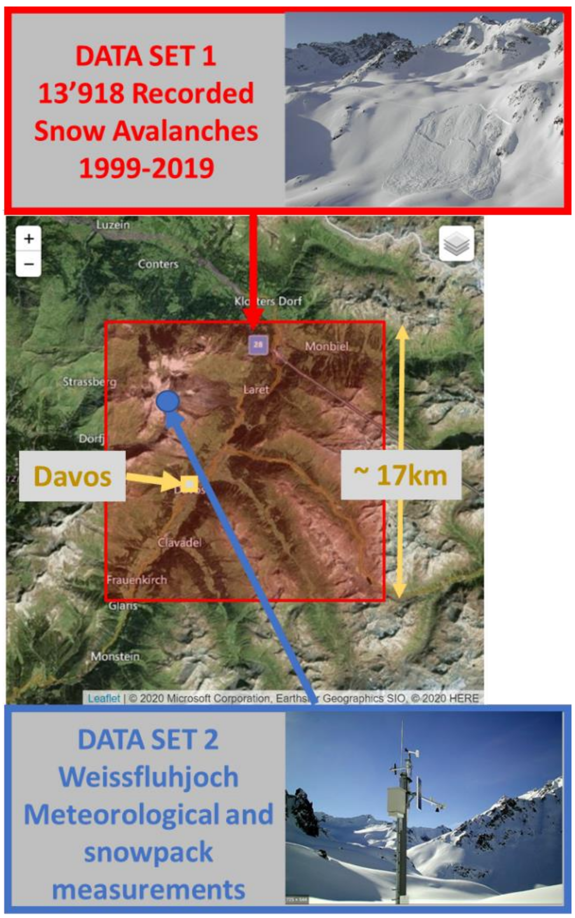


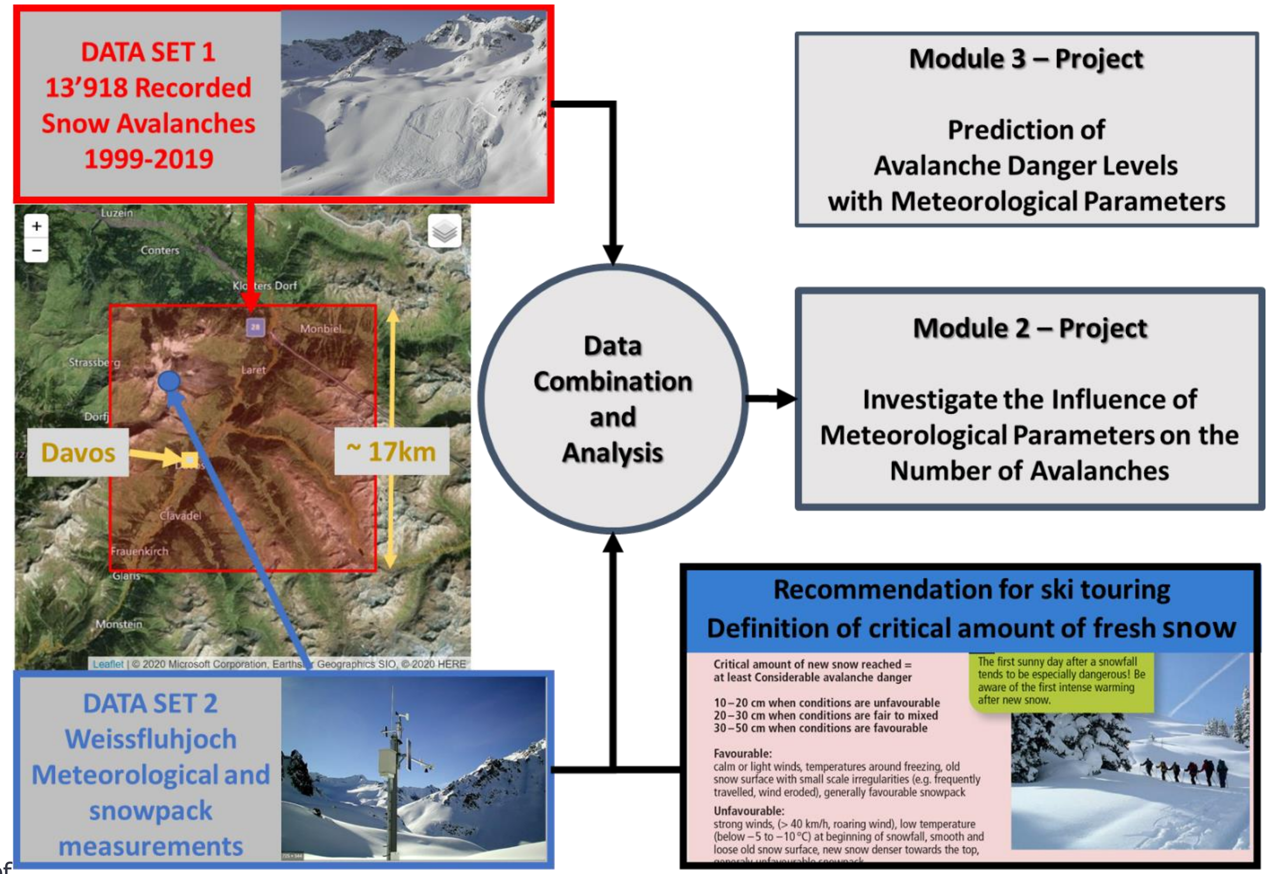
FIGURE 9.: AVALANCHE RECORDING AREA AND WEATHER STATION LOCATION

Manually recorded Variables in avalanche dataset

The variables in the avalanche dataset are recorded manually. Recording mistakes are possible and could have an influence on the data quality and on the results of this analysis.

6. Data Flow

Figure 7. below represents the overall concept of this project and the data flow



pf

FIGURE 10. DATA FLOW / PROJECT CONCEPT OVERVIEW

The two datasets were downloaded from "EnviDat", an environment data portal (www.envidat.ch) on my laptop. All the processing and analysis is done on my laptop and the outputs are:

1. The Conceptual Design Report document, with tables and figures
   https://github.com/lionelperret76/M1-M2_Project/blob/main/Lionel_Perret-CAS-ADS-M1_M2-Project_Report.pdf

2. The Modul 2 project tables and figures available in the poster file
   https://github.com/lionelperret76/M1-M2_Project/blob/main/Lionel_Perret-CAS-ADS-M2-Poster.pdf

All the other used files, datasets, and python Jupyter notebook scripts are available as well on github:
https://github.com/lionelperret76/M1-M2_Project

7. Data Model

Conceptual

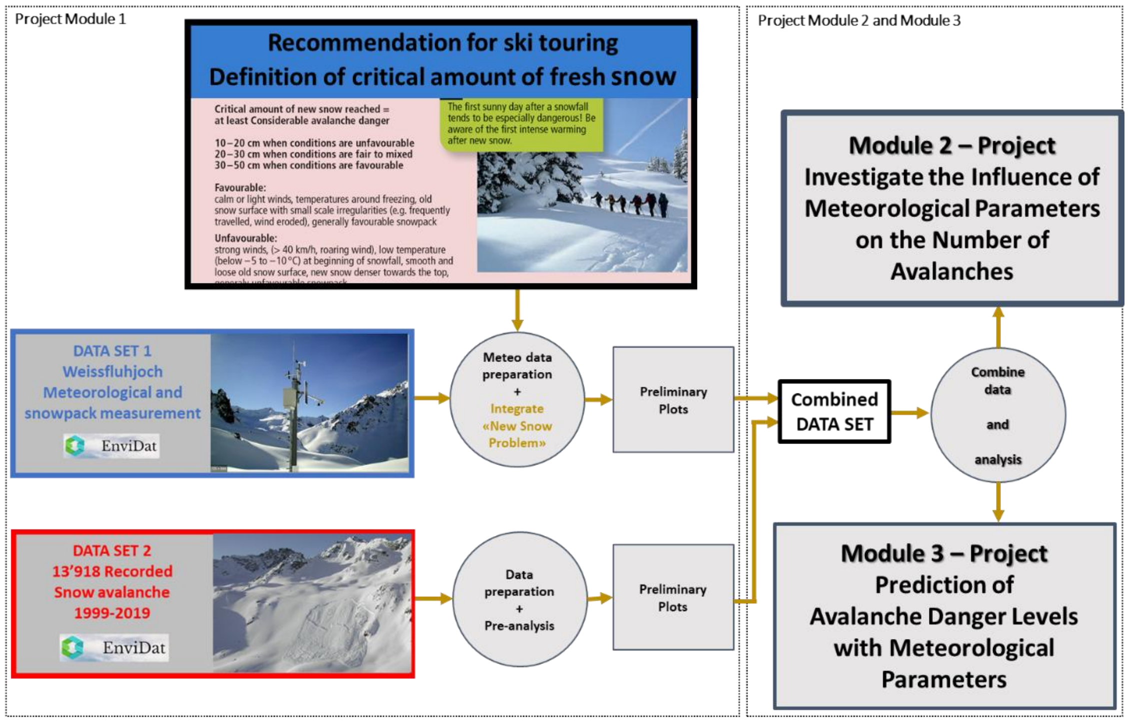


FIGURE 11.: CONCEPTUAL DATA MODEL

Logical

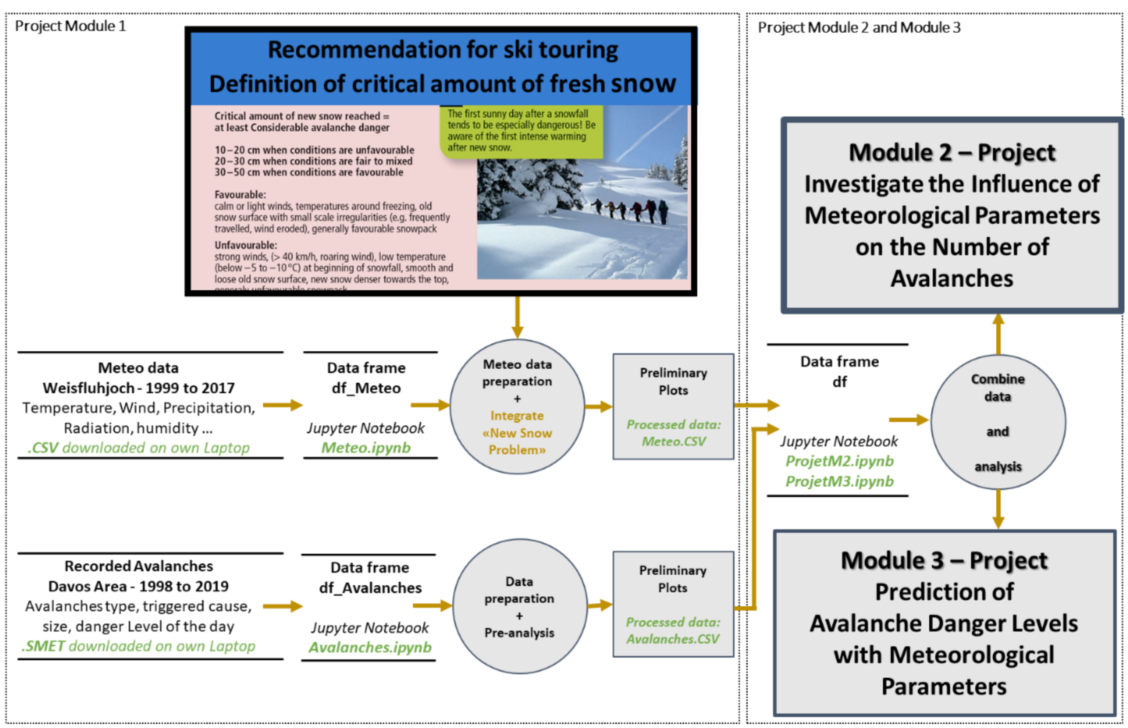


FIGURE 12.: LOGICAL DATA MODEL

Physical

There is no need on external infrastructure and all this project will be done on my personal laptop/github.

8. Risks

Recording mistakes risk

One of the risks is linked to the quality of the recorded data. Mistake could have been made in the recording of the avalanche variables (date, size, total number,) as already mentioned on the data quality chapter.
The avalanche and meteorological datasets are linked with the date parameter, so possible mistakes in the date of the occurred avalanches could have an influence on the quality of this analysis.

Analysis mistakes risk

Another risk is data analysis mistake, as there could be bugs in my python code and in the use of the different available python functions and libraries.

Data collection risk (Project Module 3)

The objective for the module 3 project is:

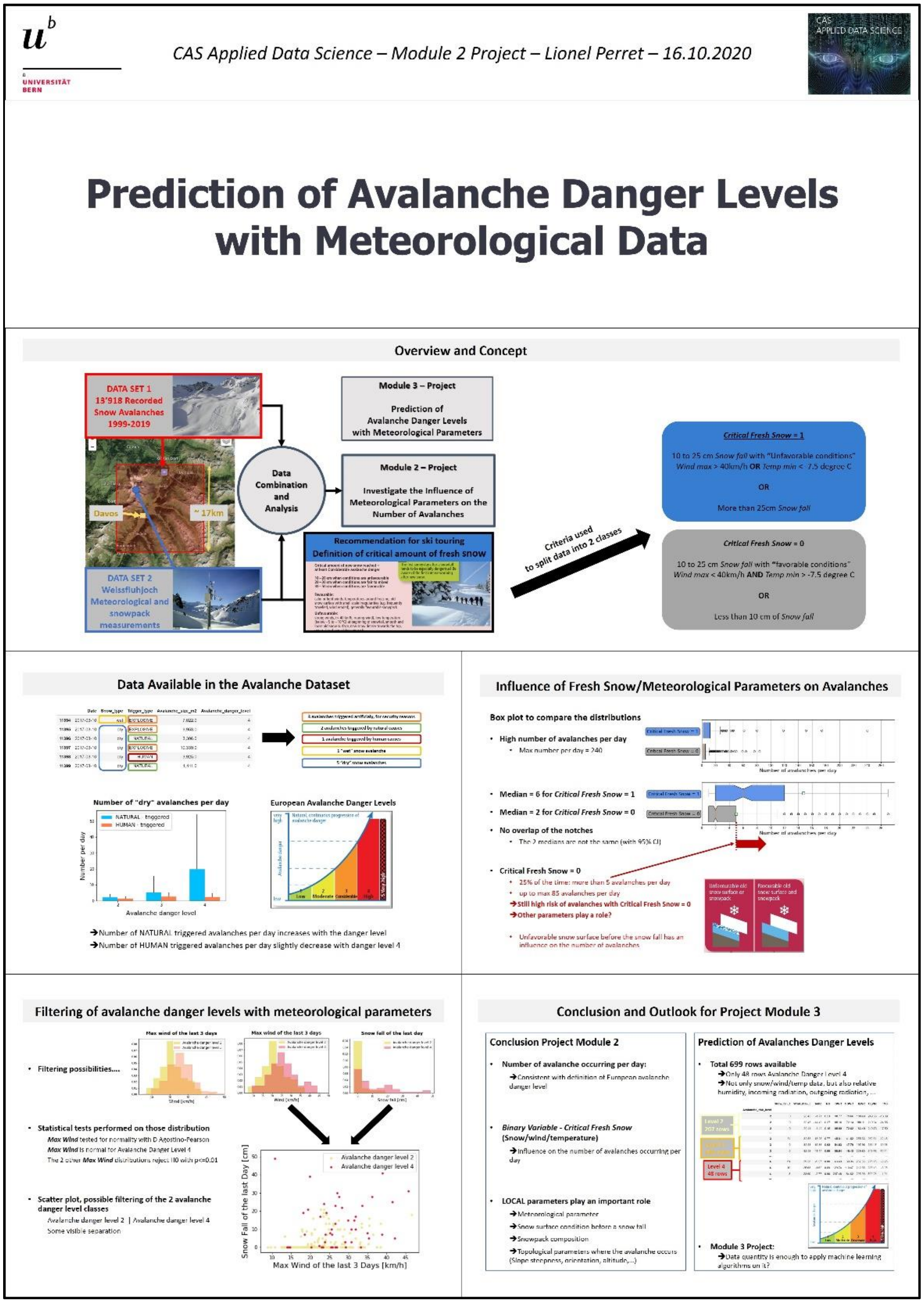
"The final goal of this project is to use and interpret the recorded meteorological data in the Weissfluhjoch defined area to predict the avalanche danger levels within this area.
To do that the model will be trained in using the two combined datasets used in module 1-2 and the model will be used to predict the real avalanche danger levels occurring in the upcoming winter 2020-2021 in the Davos area."

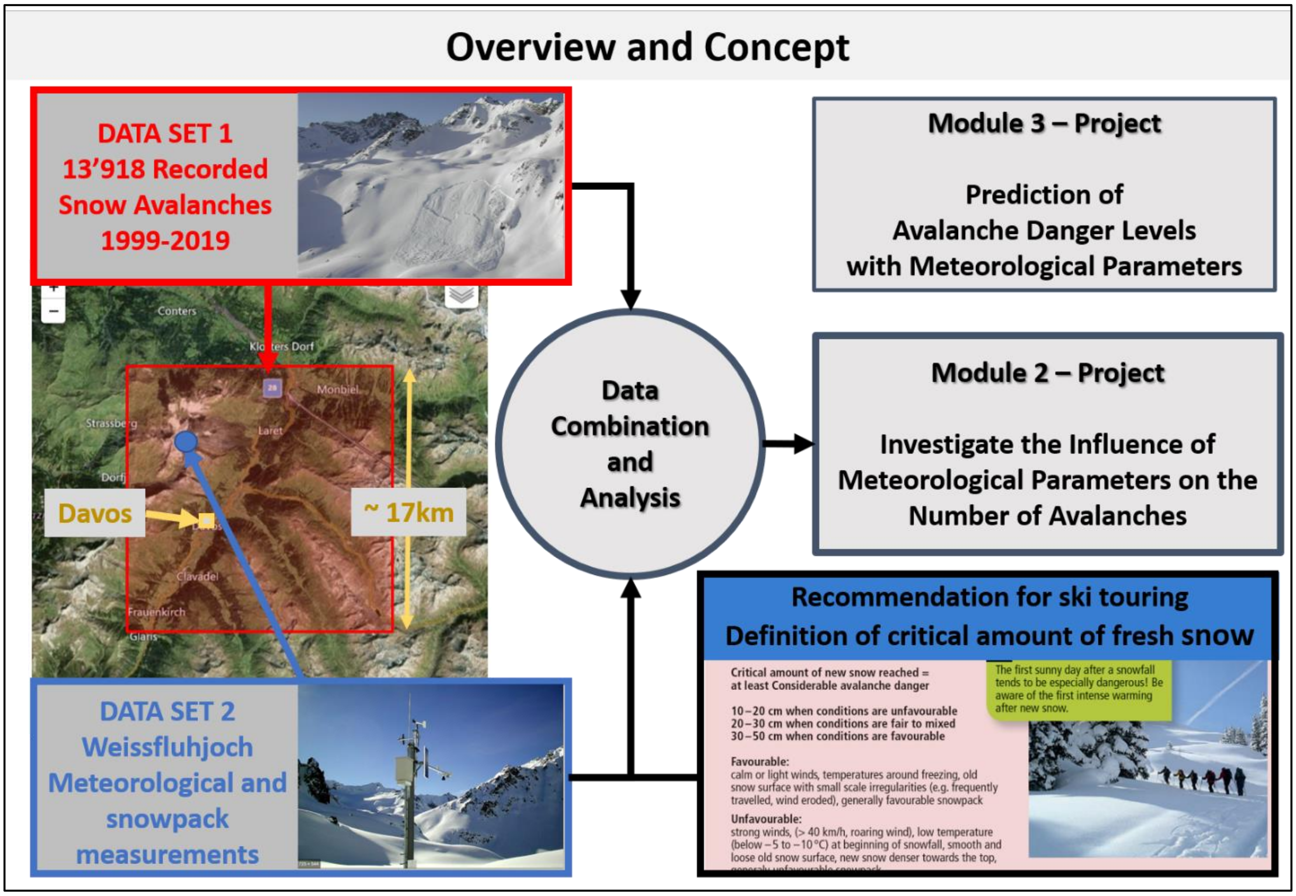
This implies that Meteorological data from the winter 2020-2021 is collected. A solution for the automatic collection of those meteorological parameters needs to be found and there is a risk that the data cannot be collected as needed.
In case the collection of the data do not succeed, the backup plan is to split the available past data in a "train" and "test" sample, so that I can still investigate how accurate is the prediction.

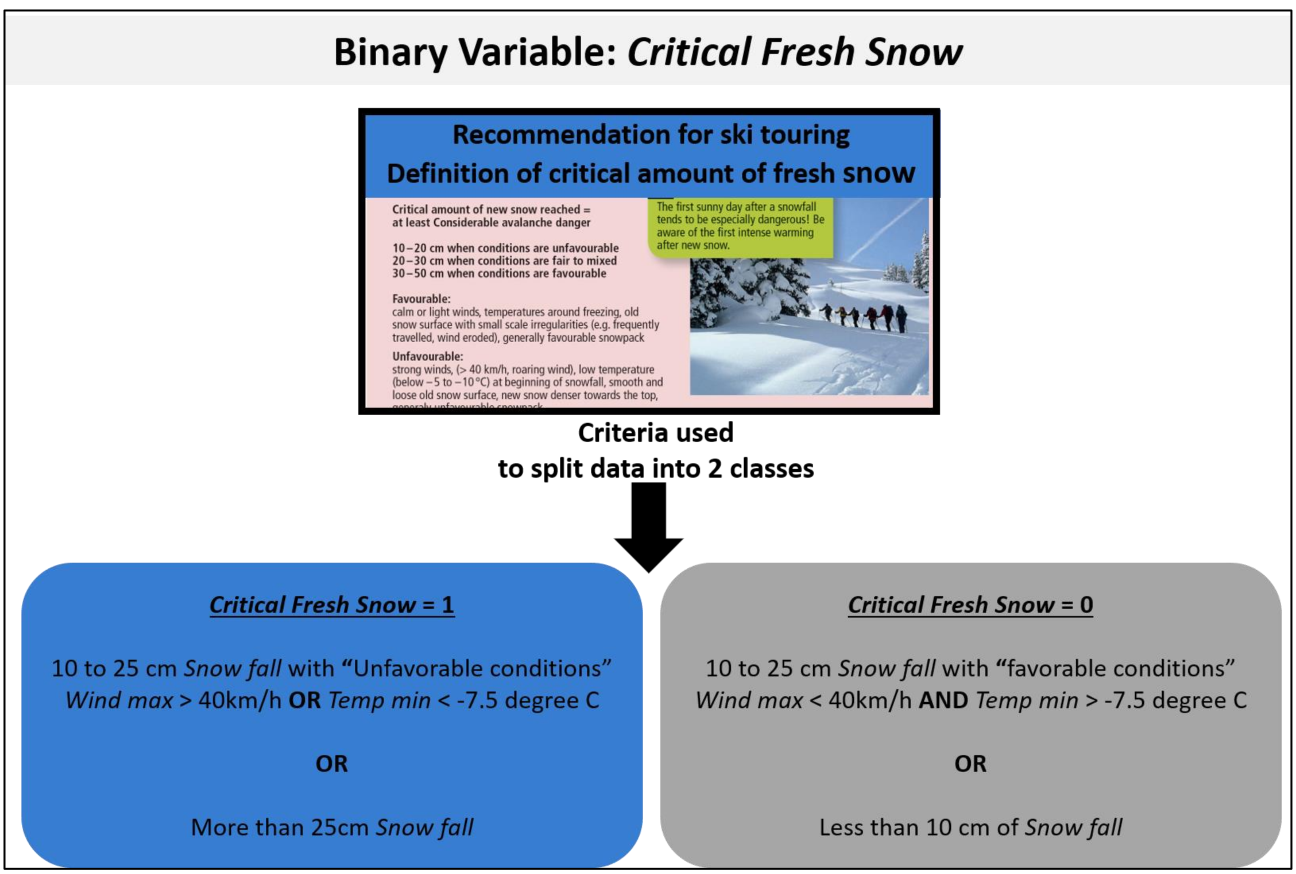
9 Preliminary Studies – Module 2

- Module 2 Jupyter Notebook is available on github:
  https://github.com/lionelperret76/M1-M2_Project/tree/main/JupyterNotebooks/ProjectM2.ipynb
  This file was used to create the different figures-tables-illustration shown in my poster.

- A1 poster (pdf file) available on github:
  https://github.com/lionelperret76/M1-M2_Project/tree/main/Lionel_Perret-CAS-ADS-M2-Poster
  The different parts of this poster are also put as printscreen in the next pages




Overview and Concept
DATA SET 1
13'918 Recorded Snow Avalanches
1999-2019
Davos
~ 17km
DATA SET 2
Weissfluhjoch
Meteorological and snowpack measurements
Data Combination and Analysis
Module 3 – Project
Prediction of Avalanche Danger Levels with Meteorological Parameters
Module 2 – Project
Investigate the Influence of Meteorological Parameters on the Number of Avalanches
Recommendation for ski touring
Definition of critical amount of fresh snow

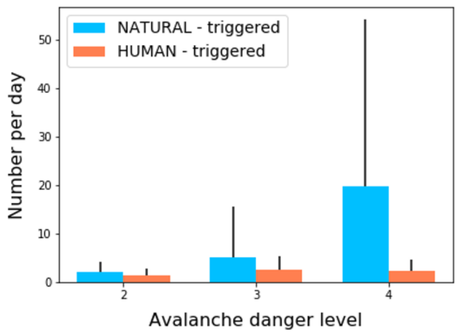

Binary Variable: Critical Fresh Snow
Recommendation for ski touring
Definition of critical amount of fresh snow
Critical amount of new snow reached = at least Considerable avalanche danger
10–20 cm when conditions are unfavourable
20–30 cm when conditions are fair to mixed
30–50 cm when conditions are favourable
Favourable:
calm or light winds, temperatures around freezing, old snow surface with small scale irregularities (e.g. frequently travelled, wind eroded), generally favourable snowpack
Unfavourable:
strong winds, (> 40 km/h, roaring wind), low temperature (below –5 to –10 °C) at beginning of snowfall, smooth and loose old snow surface, new snow denser towards the top,
The first sunny day after a snowfall tends to be especially dangerous! Be aware of the first intense warming after new snow.
Criteria used to split data into 2 classes
Critical Fresh Snow = 1
10 to 25 cm Snow fall with "Unfavorable conditions"
Wind max > 40km/h OR Temp min < -7.5 degree C
OR
More than 25cm Snow fall
Critical Fresh Snow = 0
10 to 25 cm Snow fall with "favorable conditions"
Wind max < 40km/h AND Temp min > -7.5 degree C
OR
Less than 10 cm of Snow fall

Data Available in the Avalanche Dataset

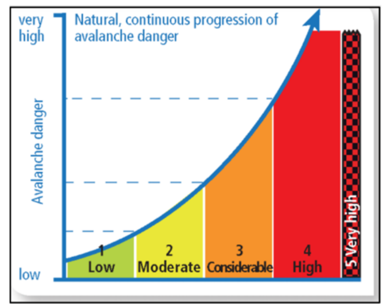
	Date	Snow_type	Trigger_type	Avalanche_size_m2	Avalanche_danger_level
11394	2017-03-10	wet	EXPLOSIVE	7,022.0	4
11395	2017-03-10	dry	EXPLOSIVE	9,953.0	4
11396	2017-03-10	dry	NATURAL	3,306.0	4
11397	2017-03-10	dry	EXPLOSIVE	10,339.0	4
11398	2017-03-10	dry	HUMAN	3,925.0	4
11399	2017-03-10	dry	NATURAL	1,411.0	4

- 3 avalanches triggered artificialy, for security reasons
- 2 avalanches triggered by natural causes
- 1 avalanche triggered by human causes
- 1 "wet" snow avalanche
- 5 "dry" snow avalanches

Number of "dry" avalanches per day



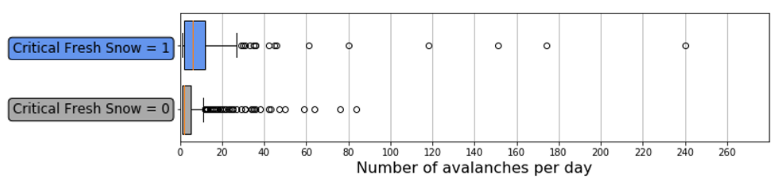
European Avalanche Danger Levels



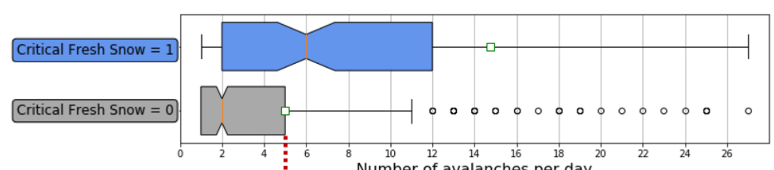
➔Number of NATURAL triggered avalanches per day increases with the danger level
➔Number of HUMAN triggered avalanches per day slightly decrease with danger level 4

Influence of Fresh Snow/Meteorological Parameters on Avalanches

Box plot to compare the distributions



- **High number of avalanches per day**
  - Max number per day = 240



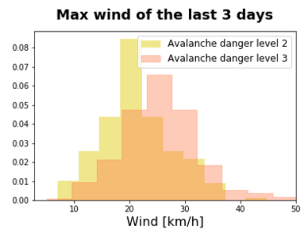
- **Median = 6 for *Critical Fresh Snow* = 1**
- **Median = 2 for *Critical Fresh Snow* = 0**
- **No overlap of the notches**
  - The 2 medians are not the same (with 95% CI)
- **Critical Fresh Snow = 0**
  - 25% of the time: more than 5 avalanches per day
  - up to max 85 avalanches per day

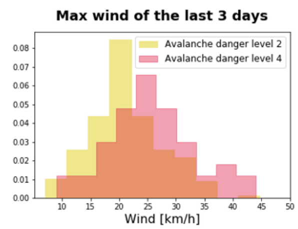
  ➔Still high risk of avalanches with Critical Fresh Snow = 0
  ➔Other parameters play a role?

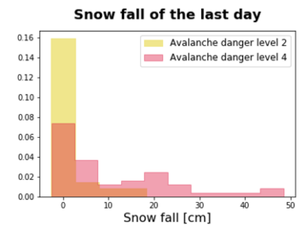
  - Unfavorable snow surface before the snow fall has an influence on the number of avalanches

Filtering of avalanche danger levels with meteorological parameters

- **Filtering possibilities….**







- **Statistical tests performed on those distribution**

  Max Wind tested for normality with D Agostino-Pearson

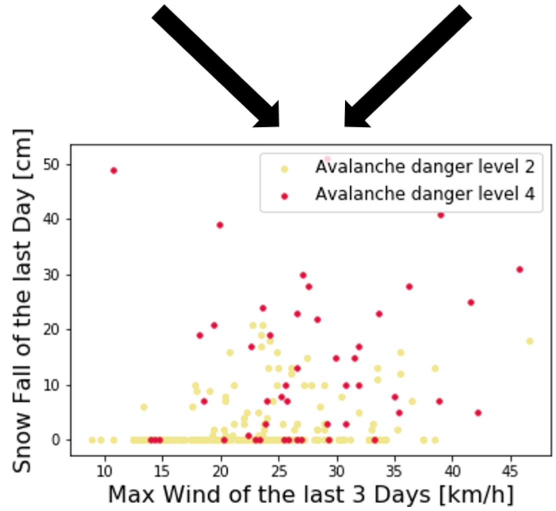
  Max Wind is normal for Avalanche Danger Level 4

  The 2 other ***Max Wind*** distributions reject H0 with p<=0.01

- **Scatter plot, possible filtering of the 2 avalanche danger level classes**

  Avalanche danger level 2 | Avalanche danger level 4

  Some visible separation



Conclusion and Outlook for Project Module 3

Conclusion Project Module 2

- **Number of avalanche occurring per day:**

  ➔Consistent with definition of European avalanche danger level

- ***Binary Variable - Critical Fresh Snow* (Snow/wind/temperature)**

  ➔Influence on the number of avalanches occurring per day

- **LOCAL parameters play an important role**

  ➔Meteorological parameter

  ➔Snow surface condition before a snow fall

  ➔Snowpack composition

  ➔Topological parameters where the avalanche occurs (Slope steepness, orientation, altitude,…)

Prediction of Avalanches Danger Levels

- **Total 699 rows available**

  ➔Only 48 rows Avalanche Danger Level 4

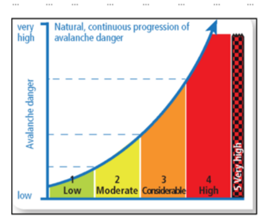
  ➔Not only snow/wind/temp data, but also relative humidity, incoming radiation, outgoing radiation, …

Level 2 207 rows

Level 3 444 rows

Level 4 48 rows

Avalanche_risk_level	Snow_fall_1	Wind_max_3	Temp	RH	ISWR	OSWR	ILWR	OLWR	TSS
2	0	28.48	-1.71	0.21	98.77	79.66	188.43	249.86	-18.09
2	0	30.46	-14.40	0.37	85.10	72.14	156.11	219.34	-24.55
2	0	20.41	-1.42	0.15	85.68	72.02	183.49	249.73	-17.50
...	...	...	...	...	...	...	...	...	...
3	31	32.58	-13.25	0.77	45.01	41.62	255.33	262.31	-12.48
3	0	32.58	-15.36	0.63	54.63	47.79	197.96	231.47	-19.58
3	0	32.58	-11.44	0.66	56.64	49.40	229.80	249.78	-15.14
...	...	...	...	...	...	...	...	...	...
4	51	29.23	-9.27	0.80	43.49	39.96	267.36	275.75	-9.25
4	10	25.63	-1.62	0.69	129.74	110.47	212.85	285.13	-8.05
4	3	23.90	-2.77	0.90	207.48	184.62	295.99	309.26	-1.91
...	...	...	...	...	...	...	...	...	...



- **Module 3 Project:**

  ➔Data quantity is enough to apply machine learning algorithms on it?

10. Conclusion

Project Module 1

After processing the 2 datasets some basic statistics like avalanches sizes was extracted from the data. From the meteorological Weissfluhjoch data, several parameters were calculated to derive to create the *Critical Fresh Snow* Binary variable to use for some part of the module 2 Project.
Python/Jupyter notebook was used to do that and is a very interesting and user-friendly tool for data collection, processing and pre-analysis.
Some recording mistakes were found in the **Date** parameter of the recorded avalanche by cross-checking with meteorological snow fall (more on how this was done is explained in comments in the Avalanches.ipynb and Meteo.ipynb files), but overall the data quality is good.

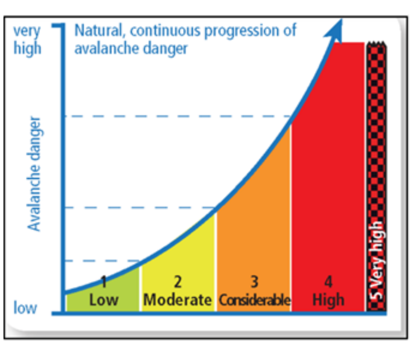
Project Module 2

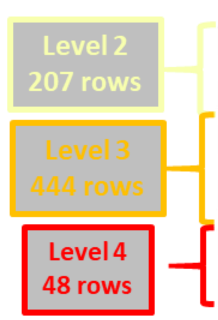
The Number of avalanches per day increases with the avalanche danger level and this is consistent with the definition of the European avalanche danger level.
The created Binary variable *Critical Fresh Snow* (with wind/temperature parameter) was used to classify the data and it gives an indication on the number of avalanches per day, but other LOCAL parameters like Snow Surface before a snow fall needs to be considered as well

Outlook Project Module 3

The aim of the Project in Module 3 is to predict Avalanche Danger Levels with Meteorological Parameters. With the combination of the data, 699 rows are available with meteorological parameters:

Is there enough data to apply machine learning algorithms on this dataset?





Avalanche_risk_level	Snow_fall_1	Wind_max_3	Temp	RH	ISWR	OSWR	ILWR	OLWR	TSS
2	0	28.48	-1.71	0.21	98.77	79.66	188.43	249.86	-18.09
2	0	30.46	-14.40	0.37	85.10	72.14	156.11	219.34	-24.55
2	0	20.41	-1.42	0.15	85.68	72.02	183.49	249.73	-17.50
...	...	...	...	...	...	...	...	...	...
3	31	32.58	-13.25	0.77	45.01	41.62	255.33	262.31	-12.48
3	0	32.58	-15.36	0.63	54.63	47.79	197.96	231.47	-19.58
3	0	32.58	-11.44	0.66	56.64	49.40	229.80	249.78	-15.14
...	...	...	...	...	...	...	...	...	...
4	51	29.23	-9.27	0.80	43.49	39.96	267.36	275.75	-9.25
4	10	25.63	-1.62	0.69	129.74	110.47	212.85	285.13	-8.05
4	3	23.90	-2.77	0.90	207.48	184.62	295.99	309.26	-1.91
...	...	...	...	...	...	...	...	...	...

Broader outlook towards "LOCAL" Avalanche Danger Level Prediction

Currently REGIONAL avalanche prediction is published by the SLF on a daily basis.
No GPS coordinates were recorded in the used Davos avalanche dataset. Having datasets of recorded avalanches with their GPS coordinates would open the possibility to link not only with meteorological parameters, but as well with topological information (steepness of the slope, altitude, orientation,…). As those topological and meteorological LOCAL parameters have a huge influence on the avalanche danger, the availability of such information, would be a first step in order to apply machine learning algorithm for a more LOCAL prediction of the avalanche danger.

Documents, Tools, Data and the WHITE RISK app **[5]** (screen shots in FIGURE 13. below), developed by the Institute for Snow and Avalanche Research, Davos (SLF) is a good first step towards this broader objective and are of great support for the people that want to stay safe in the mountain.

FIGURE 13.: SCREEN SHOTS FROM THE WHITE RISK APP

11. Acknowledgements

I would like to thank the Institute for Snow and Avalanche Research, Davos, as the information available on its website (https://www.slf.ch/en/index.html) is very accurate, well presented and interesting. This information as well as the WHITE RISK app, was not only of great help for this project, but also for acquiring knowledge to minimize the risk when going ski touring.

I would like to thank as well Fanny Viret and my classmate Alfonso Garcia Miguel, for helping me in the review of this project.

12. References

[1] Leaflet **"Caution Avalanches"**
Leaflet published by the SLF, Institute for Snow and Avalanche Research, Davos
https://www.slf.ch/en/publicationssearch/books-and-brochures.html

[2] Webpage **Avalanche danger scale, European Avalanche Danger Level**
SLF, Institute for Snow and Avalanche Research, Davos
https://www.slf.ch/en/avalanche-bulletin-and-snow-situation/about-the-avalanche-bulletin/danger-levels.html

[3] Data Set ***Snow avalanche data Davos, Switzerland, 1999-2019***
https://www.envidat.ch/#/metadata/snow-avalanche-data-davos
Authors: *Jürg Schweizer; Christoph Mitterer; Frank Techel; Andreas Stoffel; Benjamin Reuter*

[4] Data Set ***WFJ_MOD: Meteorological and snowpack measurements from Weissfluhjoch, Davos, Switzerland***
https://www.envidat.ch/#/metadata/10-16904-1
Author: *Nander Weber*

13. Table list

TABLE 1. *AVALANCHE DANGER SCALE, EUROPEAN AVALANCHE DANGER LEVELS [2]*
TABLE 2. *DATA FROM THE ORIGINAL AVALANCHE DATASET*
TABLE 3. *DESCRIPTIVE STATISTIC "DRY" AVALANCHE SIZE*
TABLE 4. *RAW DATA FROM THE METEOROLOGICAL DATA SET*
TABLE 5. *PROCESSED DATA FROM THE METEOROLOGICAL DATA SET (METEO.CSV FILE)*

14. Figure list

FIGURE 1. *CRITERIA THAT DEFINES CRITICAL FRESH SNOW CONDITIONS*
FIGURE 2. *NUMBER OF AVALANCHES WITH HUMAN/NATURAL CAUSES*
FIGURE 3. *HISTOGRAM AVALANCHE SIZE*
FIGURE4. *BOX PLOT AVALANCHE SIZE*
FIGURE 5. *SIMPLIFIED CRITERIA FOR THE CRITICAL FRESH SNOW VARIABLE*
FIGURE 6. *MEAN DAILY TEMPERATURE - WEISSFLUHJOCH WEATHER STATION*
FIGURE 7. *SUM OF SNOW FALL OF LAST 3 DAYS - WEISSFLUHJOCH WEATHER STATION*
FIGURE 8. *MEAN DAILY WIND SPEED - WEISSFLUHJOCH WEATHER STATION*
FIGURE 9. *AVALANCHE RECORDING AREA AND WEATHER STATION LOCATION*
FIGURE 10. *DATA FLOW / PROJECT CONCEPT OVERVIEW*
FIGURE 11. *CONCEPTUAL DATA MODEL*
FIGURE 12. *LOGICAL DATA MODEL*
FIGURE 13. *SCREEN SHOTS FROM THE WHITE RISK APP*