ACCIDENTS AT WORK Analysis in Italy

(2013-2018)

ABOUT THE PROJECT

The project work we are going to illustrate aims to answer the three main research questions:

1. *Which are the **REGIONS** with the highest rate of work-related accidents and fatal accidents in the workplace in Italy (taking into account the employed proportion of the population for each region)?*
2. *Which are **SECTORS** with the highest risks (adjusted proportionally as above)?*
3. *How many people **CARE** about improving job security following a serious accident? And do these mortal accidents affect workers' caution in the days following the event?*

As a data foundation for this undertaking we used three public datasets available on the Open-Data section in INAIL's website. INAIL is the Italian organization tasked with monitoring and supporting workers in the event of accidents or diseases on the job. Datasets are divided by time (2013-2017 and 2018) and by regions.

In our opinion, **the topic we faced is not taken into great consideration by our society**, which should increase awareness of the risks we face every day as workers or as citizens. For that reason we consider our work as a starting point for a **data journalism project**.

The first set dataset's title reads: "ACCIDENTS AT WORK in Italy (2013-2018)". It contains data on accidents upgraded every six months for the period from 2013 through 2017, and data upgraded monthly for the year 2018. This data can be found following these links to INAIL's website:

- o https://dati.inail.it/opendata/default/Daticadenzasemestrale/index.html
- o https://dati.inail.it/opendata/default/Daticadenzamensile/index.html

To better comprehend the information contained in the previous tables consulting the key through the following links can be helpful:

- o https://dati.inail.it/opendata/elements/Provincia
- o https://dati.inail.it/opendata/elements/LuogoNascita
- o https://dati.inail.it/opendata/elements/SettoreAttivitaEconomica

The second array of datasets is called: "EMPLOYED by sector and by region", and contains data on the number of people employed throughout the years for each Italian region. It can be consulted by following this link on INAIL's website: http://dati.istat.it/Index.aspx?DataSetCode=DCCV_OCCUPATIT1

The final dataset, named "INTEREST in JOB SAFETY", contains data pertaining to the interest from Italian citizens in fostering and supporting improved job security. Data is obtained from Google Trends: https://trends.google.it/trends/explore?date=2013-01-01%202018-12-31&geo=IT&q=%2Fm%2F035sx5.

We manually downloaded all the datasets provided by the links above, we check the consistency and we structured the steps to face individually.

SOCIAL IMPACT members
• DANILO FERRARA - data cleaning and data combining
• LEONARDO VENTURINI - data combining and data visualization
• VALERIA PICCOLI - data analysis and documentation
• CARLO CREMONA - data visualization and documentation
• FRANCESCO GHEZZE - data analysis and data visualization

STRUCTURE AND MAIN PROBLEMS OCCURRED

We divided our project in two parts: data management (**main.ipynb**) and data visualization (**results.ipynb**).

All the steps of our work are well explained in the file **main.ipynb** (see it in the root folder of the project), where it is possible to start the three main sources of our data wrangling: **data_cleaning.py**, **data_combining.py**, **data_pre_visualizing**.

In order to visualize the results of our research we prepared a specific *jupyter notebook* called **results.ipynb**.

Here there is a summary of main problems occurred:

PROBLEMS

MISSING DATA FROM 2018

After a first plot of our cleaned datasets, we discovered that data related to 2018 were incomplete (they were present only data of the first two months). We immediately wrote an email to the responsible of the opendata of INAIL which mail back to us after a couple of days. He explained to us the reason (data of the last two years are uploaded until the current month) and sended to us the complete data of 2018 😊.

DIFFERENCES IN THE NAMES OF THE REGIONS

We faced several problems in order to combine the data using as key the names of the italian regions because they were slightly different (e.g. Emilia Romanga, Emilia-Romagna). This first problem emerged in particular during the data combining between the dataset of employed and accidents. Then we incurred in the same problem while we were creating plots using the file "Italy.json" that is built with regions polygon on the map. To solve this problem we used a different key, that was made by the first 5 letters of the given regions, that were unmistakable.

DATA NORMALIZATION BETWEEN YEARS

This problem occurred during the visualization of data in the map. To show how many accidents/deaths were per region, we use different colors in the same chromatic scale to make that more understandable. The problem begun when the range of data changed every year, so it could be misunderstood. We had to standardize the range creating a personalized function because, differently from geopandas, in folium there is no function like the one we needed.

CHROMEDRIVER

To get screenshot of HTML file that we created, we used a function of selenium that make it possible. Unfortunately, Google has updated the chrome driver in 18th May, making the feature not working, because it was built for older version. We pass through this problem with a downgrade of chrome to the last version before the current.

1st RESEARCH QUESTION

Which are the regions with the highest rate of work-related accidents and fatal accidents at work in Italy (with respect to the number of employed per region)?

NOTIFICATION OF ACCIDENTS AT WORK IN THE ITALIAN REGIONS

In 2018 the reports of accidents at work presented to INAIL increased by 0.9%, probably due to the increase in the retirement age which forced thousands of old people to keep dangerous jobs that are becoming increasingly difficult to manage due to concentration,ailments and slow reflexes.

In the report presented by INAIL, Trentino Alto Adige is the region with the highest number of accidents at work reported from 2013 to 2018, followed by Emilia Romagna.

The result is given by the greater job offer, low undeclared employment rate, and a greater propensity to report compared to the rest of the Italian regions, but also and above all by the greater number of labor inspectors present in its provinces.

The sectors most at risk of accidents are agriculture and forestry, for the latter there was in fact an increase in the labor force linked to the restoration of the conditions of the Trentino forests after the wave of bad weather that hit Northern Italy in 2018.

FATAL ACCIDENTS AT WORK IN THE ITALIAN REGIONS

Across the years in Italy the rate of accidents at work has fallen by, and paradoxically the number of victims has increased according to data provided by INAIL on the basis of the complaints presented to the Institute.

Among the main risk factors is the lack of safety at work, in agriculture as well as in construction industry, and undeclared work.

By observing and analysing the information from the INAIL dataset we found a substantial high rate of fatal accidents at work in two Italian Regions in particular: Molise and Abruzzo.

The lack of safety at work is due especially to the use of old and damaged work vehicles or machinery, but also devoid of security devices and security courses for workers.

The causes of the injuries are mainly: falls from above, toxic fumes, crushing. But these are the apparent reasons: deaths often occur because the safety of workers is not considered a priority, workers are not adequately trained on the risks they run and lack controls.

To this scenario is added the precariousness of contracts and the difficulty in finding a new job, factors that make workers more and more blackmailable.

2nd RESEARCH QUESTION

Which are sectors with the highest risks (with respect to the number of employed per region)?

ACCIDENTS AT WORK PER SECTOR

The sectors most at risk include construction, agriculture, wholesale and retail trade and vehicle repair, warehousing and transport and the healthcare and social assistance. In the industrial sector, on the other hand, the manufacturing of metal products, machinery production and the food industry stand out for the rate of accidents at work.

FATAL ACCIDENT RATE PER SECTOR

According to data collected by the INAIL, as far as the sectors are concerned, fatal accidents at work are more common among construction workers, manufacturing activities, transport and storage, from wholesale and retail trade, repair of motor vehicles and motorcycles, followed by accommodation, restaurants and catering services, and the supply of water, sewage networks, waste management activities and rehabilitation.

3rd RESEARCH QUESTION

How much people care about job security when a serious accident occurs?

By looking at the Google Trends graph, we can see that the lowest levels of interest recorded over the years always coincide with the month of August.This can be explained by the fact that August is a month of vacation for most Italian companies, so it is understandable that there is a decrease in the accident rate and consequently also in the degree of interest. Another drop, although less significant than the previous one, is also recorded for the month of December.

OTHER ANALYSIS

ACCIDENTS AND DEATHS

The complaints presented for accidents at work occurred in 2014 decreased compared to 2013, while as it concerns fatal accidents at work, there was a decrease compared to the previous year.

In 2015 the decreasing trend in the number of accidents is confirmed, and there was a decrease compared to 2014. As it concerns fatal accidents at work, there was a decrease compared to the previous year, due to the increasing number of firms benefiting from the reduction in INAIL premiums for prevention merits.

In 2016 there was no significant difference compared to 2015 , while as it concerns fatal accidents at work, there was a decrease compared to the year 2015.

In 2017 there was no significant difference compared to 2016, while as it concerns fatal accidents at work, there was an increase compared to the year 2016. In 2017, in fact, 13 multiple fatal accidents occurred in January in Abruzzo.

We can notice that the trend of the accidents is a downtrend, from 2013 to 2018 the number of accidents has decreased, this due to factors as technological development and safety courses. However, as it is possible to notice from the first graph there was an increasing for the fatal accidents in 2018.

The expansion of the tertiary sector is one of the most significant causes in determining the decrease in the number of injuries, as it directs the workforce towards less risky occupations.

MALE AND FEMALE

On the basis of the collected results it appears that the percentage of death at work by sex has remained constant over the years, around 9% for women and around 91% for men. The reason behind this great disparity is that there are very few women hired in the sectors most at risk.

On the other hand, the percentage of non-fatal accidents at work is around 35% for women and round 65% for men. The employment status of women, mainly engaged in administrative activities of the tertiary sector or with clerical or managerial tasks in the most dangerous sectors such as industry, means that the risk of male workers is much higher than that of female workers.

AGES

As regards the analysis of accidents at work by age group, the data show that the age groups most affected are those from 45 to 54 years old workers,with an average of about 170000 injuries per year, followed by age groups from 35 to 44 years old, for which the number of accidents fell constantly: from around 170000 accidents in 2013 to around 130000 in 2018.

With regard to the age group between 25 and 34 years old, the number of accidents has steadily decreased to reach an almost unvaried number of 110000 accidents per year, while as it concerns the age group between 15 and 24 years old, there has been a constant increase in the number of injuries up to a level of 110000 injuries in the year 2018.

The data clearly show that if, on one hand, young people pay the price of precariousness and uncertainty, on the other hand older workers find themselves having to put up with the weight of often exhausting working conditions.

As regards the analysis of fatal accidents at work by age group, the data show that the age groups most affected are those from 55 to 64 years old workers, followed by age groups from 45 to 54 years old, for which the number of accidents fell constantly.

ITALIAN AND FOREIGNERS

By analyzing the data collected, it appears that the percentage of deaths at work by nationality also remained constant over the years: about 85% of deaths at work involved Italian citizens, while the remaining 15% involved foreign workers.

As it concerns non-fatal accidents at work by nationality it appears that about 64% of accidents at work involved italian citizens, while the remaining 36% involved foreign workers.