

Evolution of Total Road Accidents in Italy (2001-2018)

Name: Davina David

Matricola: 946851

Email: d.david2@campus.unimib.it

Research Questions

Primary Research Question:

How did road accidents and vehicles involvement evolve in Italy from 2001 to 2018?

Supporting Sub-Questions

- 1) Which Year recorded the most accident count?
- 2) Which regions recorded the highest accident totals per year?
- 3) Which macro-areas (North, Center, South) show the highest risk over these years?
- 4) Which vehicle contributed most over these years?

About the Data

- Dataset: Road accidents in Italy by vehicle type and region (2001–2018)
- A new dataset is extracted from original dataset for aggregated regions
- Source: MIT Open Data (Italy Ministry of Infrastructure & Transport)
- Format: CSV/XLS
- Coverage: 20 Italian regions, 18 years
- License: CC-BY-4.0 (free for reuse with attribution)
- Data Quality Considerations:
 - i. Dataset is aggregated at regional level
 - ii. Some years show unusual spikes (e.g., vehicle involvement).
 - iii. Requires cleaning: renaming headers, ensuring numeric types, replacing null values to 0 (for calculation accuracy), dataset translation to English

Methodology

- **Data Collection:**

Downloaded dataset manually from MIT Open Data portal

Created a separate dataset for aggregated region from original dataset for Macro-Regions.

- **Data Cleaning & Processing:**

- Cleaned and formatted columns in Power BI and Excel
- Translated original dataset to English
- Standardized region names by separating macro-regions
- Converted accident counts of each vehicles to numeric values
- Replaced missing rows to '0' for numeric values and 'NA' for string values for data accuracy

- **Analysis Techniques:**

- Time-series trend analysis
- Regional aggregation
- Comparative analysis (accident of regions over years)

Methodology

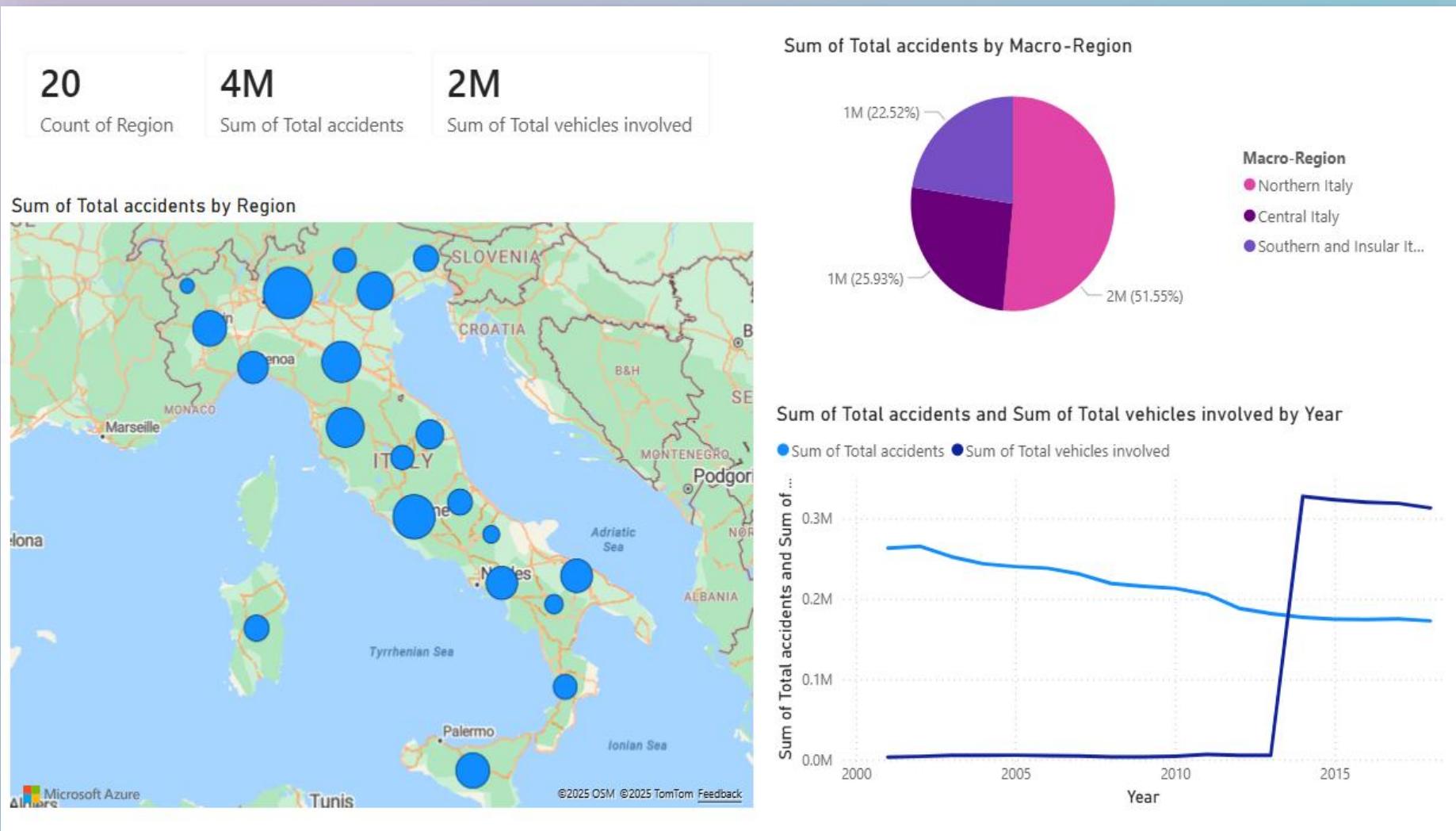
- **Tools Used:**

- Power BI (primary for cleaning + visualization)
- ChatGPT: guidance, explanation writing

- **Reproducibility:**

- The road accident dataset was obtained from MIT Open Data under the CC-BY-4.0 license and downloaded in CSV format. The analysis uses national-level aggregated files covering the period 2001–2018.
- The CSV file was imported into Power BI Desktop, where data preparation steps included renaming columns, converting variables to appropriate data types, handling missing values by replacing null accident counts with zero, and aggregating provinces into macro-regions (North, Center, South).
- All analyses and visualizations were created in Power BI using line charts, pie charts, and stacked bar charts to address the defined research questions. The final dashboard and report were then published for presentation and review.

Road Accident Evolution (2001-2018)



Dashboard Summary: Vehicle Accidents and Involved Vehicles in Italy

- **Regions Overview:**
 - **20 Regions** displayed on the map.
 - The size of the circles on the map indicates the number of accidents in each region.
- **Total Accident Data:**
 - **4 million total accidents** recorded.
 - **2 million total vehicles involved** across all regions.
- **Regional Distribution of Accidents:**
 - **51.55% of accidents** occur in **Southern and Insular Italy**.
 - **22.52% of accidents** are from **Northern Italy**.
 - **25.93% of accidents** are from **Central Italy**.
- **Trends Over Time:**
 - A **steady decline** in the number of accidents and vehicles involved from 2000 to 2015 (line chart).
 - A **sharp spike in 2015** for total vehicles involved, indicating a significant event or change in data reporting.

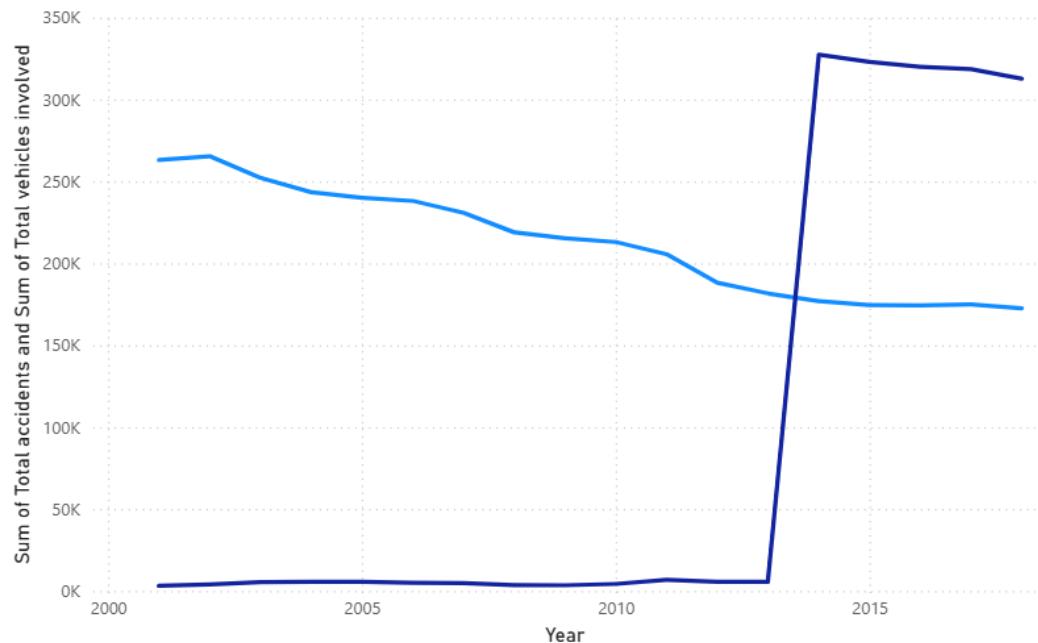
Q. *How did road accidents and vehicles involvement evolve in Italy from 2001 to 2018?*

- From 2001 to 2018, road accidents in Italy show a clear and consistent downward trend, declining from over 260,000 accidents per year in the early 2000s to around 170,000 by the late 2010s. This gradual reduction suggests improvements in road safety measures, vehicle technology, and traffic regulation over time.
- In contrast, the number of vehicles involved in accidents follows a different pattern. Until around 2012, vehicle involvement remains relatively low and stable. However, a sharp increase occurs between 2013 and 2014, after which the values stabilize at a much higher level, remaining above 300,000 vehicles involved per year.
- This abrupt change indicates a structural discontinuity in the data, likely caused by a modification in reporting practices, data definitions, or accident-recording methodology rather than a sudden real-world surge in accidents.
- Overall, while accident frequency steadily decreases, the vehicle involvement data suggests a change in measurement rather than behavior, highlighting an important data quality consideration when interpreting post-2013 trends

Accident Evolution per Year

Sum of Total accidents and Sum of Total vehicles involved by Year

● Sum of Total accidents ● Sum of Total vehicles involved



Year	Sum of Total accidents	Sum of Total vehicles involved
2002	265402	4102
2003	263100	3126
2004	252271	5441
2005	243490	5557
2006	240011	5683
2007	238124	4941
2008	230871	4733
2009	218963	3668
2010	215405	3522
2011	212997	4315
2012	205638	6776
2013	188228	5592
2014	181660	5550
2015	177031	327398
2016	174933	318629
2017	174539	322981
2018	174348	319993
Total	3829564	1664839

1. Which Year recorded the most accident count?

- The line chart indicates that **2002 recorded the highest number of road accidents**, with **over 265,000 incidents**, making it the peak year in the observed period. Following 2002, accident counts show a **continuous and sustained decline**, with particularly noticeable reductions after 2012.
- Although the line chart displays a **sharp spike in the number of vehicles involved in 2015**, this does **not correspond to a peak in accident counts**. Instead, accident totals continue to decrease during this period. This divergence further supports the interpretation that the 2015 spike reflects a **change in vehicle reporting or classification**, rather than an actual increase in accident frequency.
- Therefore, **2002 remains the year with the highest accident count**, while later anomalies in vehicle involvement should be interpreted cautiously due to likely data reporting effects.

Sum of Accidents for each Region per Annum

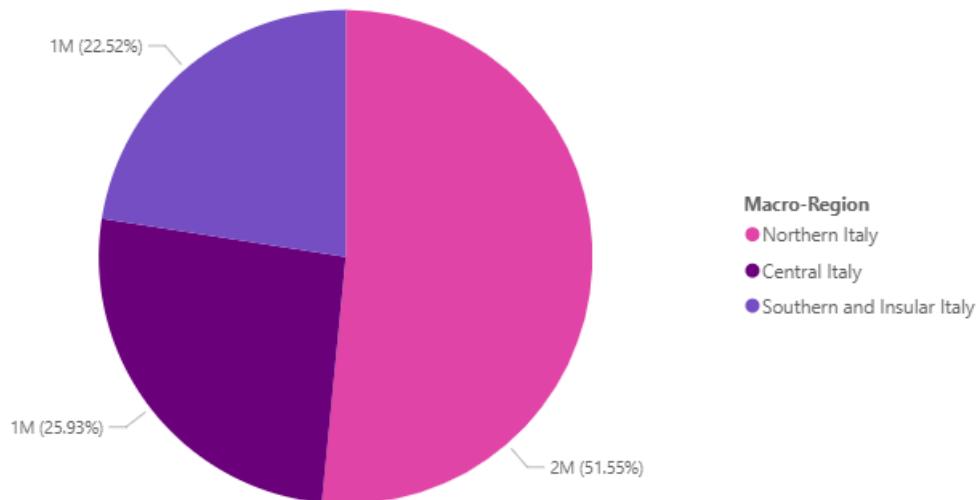
Region	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Total
Lombardia	54071	54024	51101	48627	46654	46173	44688	41827	40100	39322	37130	35612	33997	33176	32774	32785	32552	32553	737166
Lazio	33024	32569	31366	32959	32618	31344	29761	27735	28186	27810	26892	23745	22168	20589	20227	20227	19590	18613	479423
Emilia Romagna	27457	27272	26454	25894	24250	23950	23074	21744	20411	20153	20415	18321	18136	17455	17385	17406	17362	16597	383736
Toscana	22445	22240	22798	21203	21133	20826	20209	18803	18362	18865	18672	17077	16231	16654	15863	15863	16099	15823	339166
Veneto	21790	20745	19213	18895	18378	19261	18378	16744	15643	15651	15564	14365	13794	13958	13867	14034	13844	14106	298230
Piemonte	16953	17994	16111	15553	15126	14871	14643	13152	13742	13580	13254	12175	11259	11445	11134	10905	10823	10832	243552
Sicilia	15389	15805	14747	13813	14412	14203	14173	14347	14044	14255	13283	11790	11823	11366	10864	10864	11056	11019	237253
Puglia	11188	11550	11029	10526	11235	11583	11776	12024	12812	12479	12101	10287	10202	9499	9524	9524	9786	9693	196818
Campania	10309	12232	11386	10220	11508	10968	11278	11529	11745	11129	10225	9698	9103	9182	9111	9111	9922	9721	188377
Liguria	10402	10555	10021	9723	9877	10085	9987	9428	9654	9702	9292	8769	8773	8387	8415	8282	8680	8286	168318
Marche	8316	8612	8405	7597	7213	7577	7149	6919	6617	6728	6535	5482	5549	5422	5333	5333	5484	5216	119487
Friuli Venezia Giulia	5925	5899	5563	5303	5015	5065	5022	4771	4494	3933	3604	3540	3304	3316	3538	3455	3468	3351	78566
Sardegna	5042	5332	4894	5200	5089	5034	4481	4408	4665	4206	3785	3472	3664	3492	3537	3537	3425	3461	76724
Abruzzo	5574	5495	5286	4977	4814	4665	4253	3981	3853	4099	4058	3671	3603	3429	3217	3217	2946	3145	74283
Calabria	4388	4705	4420	4053	3921	3717	3526	3354	3457	3378	2989	2772	2773	2659	2733	2733	2910	2929	61417
Trentino Alto Adige	4255	4009	3572	3405	3484	3350	3124	3053	2715	2620	2991	3264	3169	3002	3052	3105	3011	3099	58280
Umbria	4181	4116	3861	3721	3464	3581	3573	3312	3074	2913	2856	2363	2402	2258	2285	2285	2361	2385	54991
Basilicata	911	918	888	835	889	921	900	954	942	1147	1054	949	888	936	936	936	848	979	16831
Molise	1033	877	743	568	552	557	512	577	530	657	639	581	507	511	461	461	510	478	10754
Valle d'Aosta	447	453	413	418	379	393	364	301	359	370	299	295	315	295	283	285	256	267	6192
Total	263100	265402	252271	243490	240011	238124	230871	218963	215405	212997	205638	188228	181660	177031	174539	174348	174933	172553	3829564

2. Which regions recorded the highest accident totals per year?

- **Highest Total:** The highest total of accidents is recorded in **Lombardia** with 737,166 accidents over the years.
- **Lowest Total:** The lowest total is observed in **Valle d'Aosta**, with only 6,192 accidents during the same period.
- **Decreasing Trend:** Most regions show a steady decrease in accidents from 2001 to 2018, similar to the overall national trend.
- **Extreme Year for Lombardia:** Lombardia saw the highest number of accidents in **2001** with **54,071**, and the number decreased over the years, reaching **32,553** by 2018.
- **Regional Comparison:** **Lazio**, **Emilia Romagna**, and **Toscana** also have relatively high numbers, with totals of **479,423**, **383,736**, and **339,166** respectively.
- **Steep Decline in Many Regions:** **Sicilia** and **Puglia** saw notable decreases in the number of accidents, with numbers dropping from **15,389** in 2001 to **9,768** in 2018 for Sicilia and from **11,150** to **7,411** for Puglia.
- **Lower Numbers in Smaller Regions:** Regions like **Molise** and **Valle d'Aosta** consistently have much lower accident totals, reflecting the smaller population and fewer incidents.

Accident Distribution Across Italian Macro-Regions

Sum of Total accidents by Macro-Region



Macro-Region
● Northern Italy
● Central Italy
● Southern and Insular Italy

Macro-Region	Sum of Total accidents
Northern Italy	1974040
Central Italy	993067
Southern and Insular Italy	862457
Total	3829564

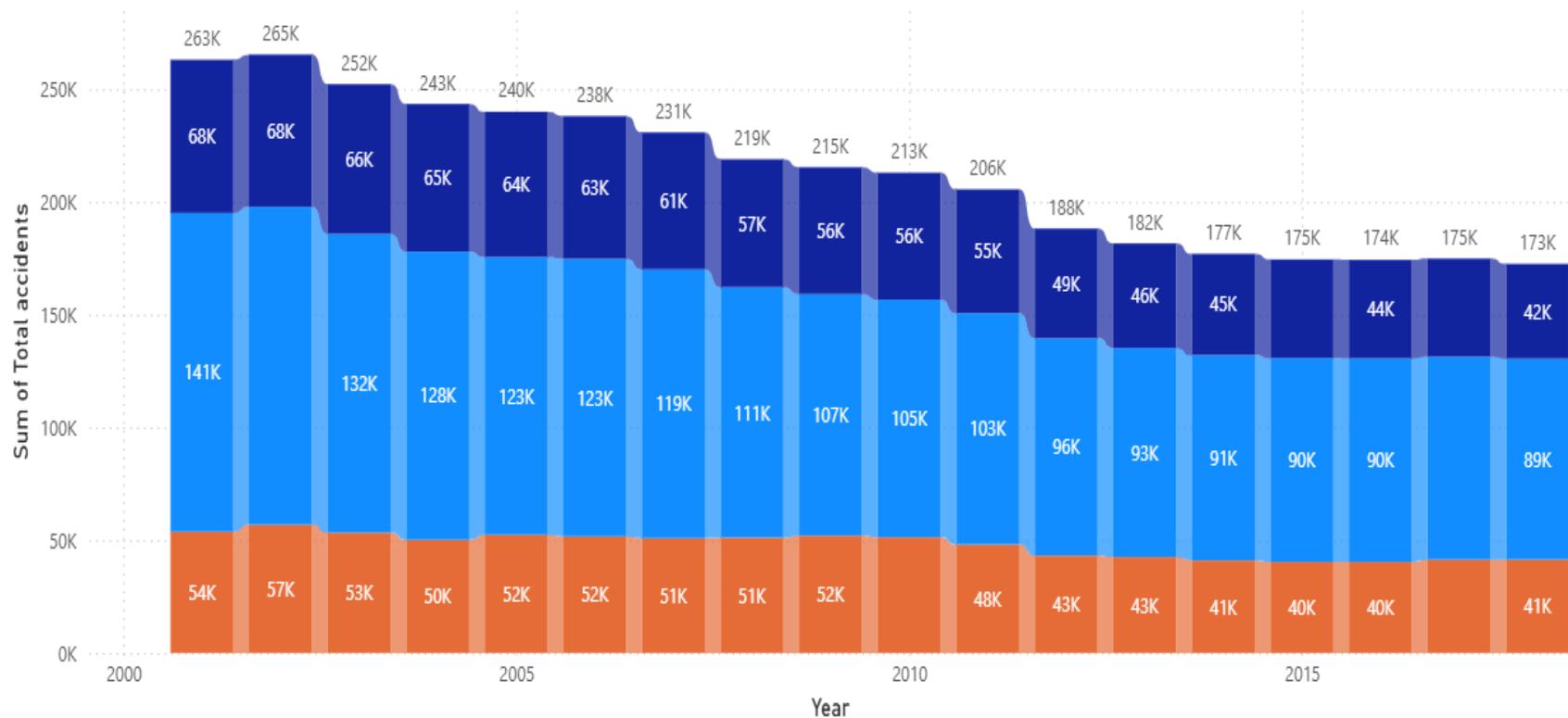
3. Which macro-areas show the highest risk over these years?

- **Northern Italy** accounts for the largest share of accidents overall, with more than half of the total accidents.
- **Southern and Insular Italy** contributes less than **Central Italy**, but still represents a substantial portion of the total accidents.
- The chart highlights the geographic distribution of accidents, emphasizing the dominance of the Northern regions in total accident counts.

Accidents in Macro-Regions per Annum

Sum of Total accidents by Year and Macro-Region

Macro-Region ● Southern and Insular Italy ● Northern Italy ● Central Italy



Accidents Contribution by Macro-Regions for each year

- **Trends Over Time:**

Accidents in Italy decreased overall from 263K in 2000 to 173K in 2015.

Northern Italy saw the highest numbers, dropping from 141K to 89K, while Southern and Insular Italy fell from 54K to 41K, and Central Italy from 68K to 42K, remaining the lowest.

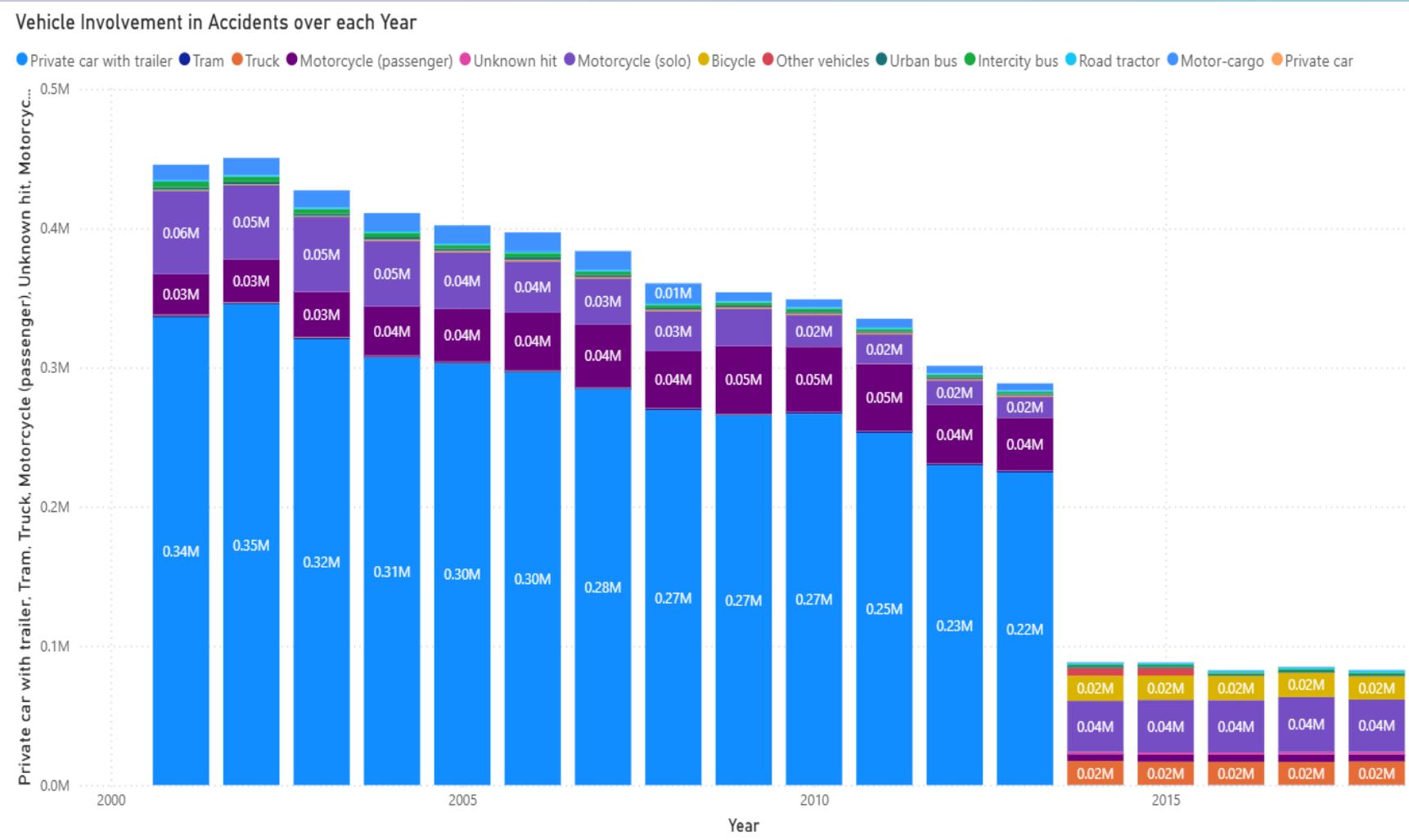
- **Accident Distribution:**

Northern Italy consistently contributed the largest share, though gradually declining. Southern and Insular Italy showed more variability but a smaller share, while Central Italy maintained the lowest counts.

- **Possible Reasons for Trends:**

The overall decline may result from improved road safety measures, stricter traffic laws, better driver awareness, and enhanced vehicle safety. Year-to-year variations could be influenced by economic factors, weather conditions, or regional infrastructure changes.

Vehicle contribution in Accident for Each Year



4. Which Vehicle Types Contribute Most to Road Accidents?

- **Trends over Time:**
 - **Private cars with trailers** consistently have the highest accident involvement, with a slight decrease over time from **0.34M** in 2000 to **0.22M** in 2015.
 - **Motorcycles** (both solo and passenger) show notable involvement, with **motorcycles (passenger)** having steady numbers and **motorcycles (solo)** showing more fluctuation over time.
 - **Other vehicles** and **bicycles** remain relatively stable but with a slight decrease over the years.
 - **Trams, buses**, and other larger vehicles tend to have much lower accident counts, contributing smaller proportions to the overall data.
- **General Trend:** Overall, the total number of accidents involving each vehicle type decreases over time, particularly with private cars. Accidents involving smaller vehicles like motorcycles and bicycles are still significant, but their share is smaller than that of private cars.
- **Private cars, especially those with trailers, contribute the most to road accidents in Italy.**

License

- 'Road Accidents in Italy Analysis (2001-2018)' dataset is licensed under CC-BY 4.0
- Dataset:

<https://dati.mit.gov.it/catalog/dataset/incidenti-per-tipo-di-veicoli-coinvolti-e-regione-anni-2001-2018>

- Power BI Visualization (Published)

<https://app.powerbi.com/groups/me/reports/906fd221-c78b-419b-aae5-dc1e8b2fd97e/3bbeef798d3d83b80837?experience=power-bi>

- Source code Access:

<https://github.com/DavinaDavid/VisualizationOfRoadAccidents>