

交通事故本票の発生年限定と暦日数調整による集計比較

研究報告書 使用資料2024年版本票

要旨

警察庁の2024年版本票290,895行から、発生年が2024年の282,376行を選んだ。月別件数の最大は10月の25,466件、暦日あたりの最大は11月の834.4件だった。時間帯別では17時台が25,593件で最大だった。件数と時間帯別の交通曝露あたり事故率は区別し、運転リスクは推定しなかった。

1. 資料の範囲と定義

2024年版の本票CSV、ファイル定義書、コード表を保存した。本票の重複を確認するキーは都道府県コード、警察署等コード、本票番号の組合せである。キーの重複はなかった。補充票や高速票を結合せず、本票を一度ずつ数えた。

収録された発生年は2024年282,376行、2023年8,477行、2022年37行、2021年3行、2020年1行、2019年1行だった。主集計は発生年2024年の行だけを使う。除外した8,519行は全体の2.93%である。資料の年版と実際の発生年を同一視しない。

事故内容の「死亡」は事故発生から24時間以内の死者を伴う区分で、「負傷」と分かれる。事故件数、死亡事故件数、死者数は異なる単位となる。ここでは主に本票件数を扱い、死者数を事故件数へ加算しない。

2. 前処理と計算

発生年月日を実在する日付として解析し、発生時は0～23時、事故内容は定義された2区分であることを確認した。使用した年月日、時刻、事故区分に欠測はなかった。コード表の曜日は日曜1から土曜7で、日付から再計算した曜日と全行一致した。

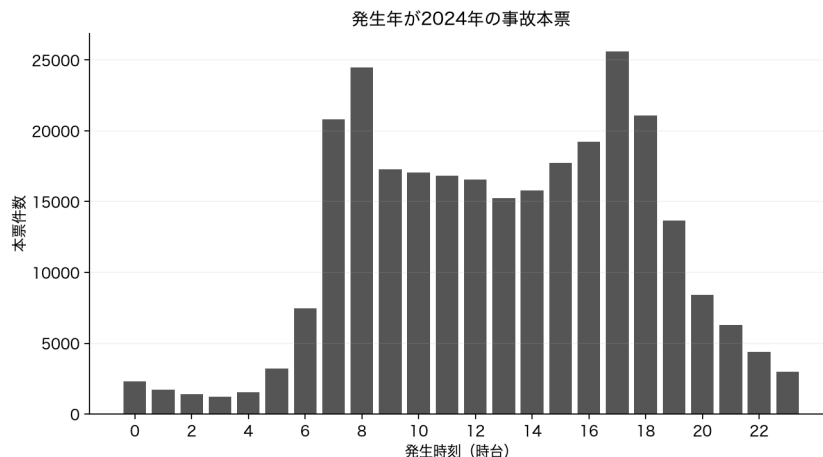
月別件数を各月の暦日数で割った。2024年2月は29日とし、年間366日とした。月～金と土日は暦の曜日による分類で、祝日を別扱いにしていない。対応する暦日数は262日と104日である。時間帯では各時台の件数と全体構成比を計算した。事故のない走行の分母は補完していない。

感度確認として発生年を限定しない全ファイルでも最多時間帯を計算した。ファイル版を変えた比較や翌年以降の遅延収録は追跡していない。

3. 結果

比較項目	結果	分母
件数最大の10月	25,466件	月全体
10月の暦日あたり件数	821.5件	31日
暦日あたり最大の11月	834.4件	30日
月～金の暦日あたり件数	822.4件	262日
土日の暦日あたり件数	643.3件	104日

17時台は25,593件、全対象の9.06%だった。全ファイルでも17時台が最多となり、今回の発生年限定によって最多時間帯は変わらなかった。発生年2024年の死亡事故区分の本票は2,591件、対象内の0.918%だった。この割合は記録された事故のうちの死亡事故割合であり、運転時の死亡確率ではない。



時間帯別の本票件数

4. 解釈と制約

暦日数による補正は、月の長さの違いだけを調整する。日別走行距離、交通量、天候、車種、道路条件、運転者構成は調整していない。時間帯別の件数の違いも交通曝露の違いを含むため、最多時間帯を最も危険な時間と呼ばない。

公開データは公表時点の情報で、別の統計表の訂正を反映しない場合がある。発生年限定後の282,376件は「2024年の公式年間発生件数」という名称では掲載しない。報告漏れや後年収録の影響は未確認である。多数の行があることだけを理由に有意差検定を加えず、事故のない観測を含む標本設計が必要な回帰も見送った。

企業への適用では運行台帳の日時、走行距離、走行時間、運行回数、車種、道路区分と対応させる。事故の報告基準やヒヤリハットの報告習慣が部署ごとに違う場合も確認する。公開時は細かな位置や個々の本票番号を出さず、必要な集計だけを使う。

5. 再現と利用条件

analyze.pyの集計に対し、verify.pyは標準CSV読込とカウンターで全290,895行を読み直し、対象件数、月別・時間帯別の全セルを照合した。monthly.csv、hourly.csv、weekday.csvの保存内容は集計値のみである。取得日時・URL・SHA-256はraw/downloads.jsonに記録した。

出典：警察庁[交通事故統計情報のオープンデータ・2024年](#)を加工して作成。[利用規約](#)の公共データ利用規約第1.0版を確認。2026年9月7日確認。警察庁の作成・承認ではなく、査読は受けていない。