|  |  |  |  |
| --- | --- | --- | --- |
| train\_kaggle  (WG 1-5)  WG 6  Variables:  Id, Datum, WG, Umsatz | 1.7.2013  1.7.2013 | 31.7.2018  27.12.2017 | > miss\_var\_summary(d1)  # A tibble: 4 × 3  variable n\_miss pct\_miss  *<chr>* *<int>* *<dbl>*  1 id 0 0  2 Datum 0 0  3 Warengruppe 0 0  4 Umsatz 0 0 |
| wetter  Variables:  Wettercode. Bewoelkung, Datum, Temp, Wind | 1.1.2012 | 1.8.2019 | > miss\_var\_summary(d3)  # A tibble: 5 × 3  variable n\_miss pct\_miss  *<chr>* *<int>* *<dbl>*  1 Wettercode 669 25.7  2 Bewoelkung 10 0.384  3 Datum 0 0  4 Temperatur 0 0  5 Windgeschwindigkeit 0 0 |
| kiwo  Variables (Boolean):  Datum, KW | 16.6.2012  22.6.2013  21.6.2014  20.6.2015  18.6.2016  17.6.2017  17.6.2018  22.6.2019 | 24.6.2012  30.6.2013  29.6.2014  28.6.2015  26.6.2016  25.6.2017  24.6.2018  30.6.2019 | > miss\_var\_summary(d2)  # A tibble: 2 × 3  variable n\_miss pct\_miss  *<chr>* *<int>* *<dbl>*  1 Datum 0 0  2 KielerWoche 0 0 |
| schulferien  Variables (Boolean):  Datum, Schulferien | 1.7.2013 | 1.8.2019 | > miss\_var\_summary(d4)  # A tibble: 2 × 3  variable n\_miss pct\_miss  *<chr>* *<int>* *<dbl>*  1 Datum 0 0  2 Schulferien 0 0 |
| feiertage\_compact  Variables (Boolean):  Datum, Feiertag | 1.7.2013 | 30.7.2019  Anpassen an train\_kaggle?  Oder wetter? | > miss\_var\_summary(d5)  # A tibble: 2 × 3  variable n\_miss pct\_miss  *<chr>* *<int>* *<dbl>*  1 Datum 0 0  2 Feiertag 0 0 |

After tibble and adding weekend/season

> miss\_var\_summary(data)

# A tibble: 12 × 3

variable n\_miss pct\_miss

*<chr>* *<int>* *<dbl>*

1 KielerWoche 9882 97.5

2 Schulferien 7861 77.6

3 Wettercode 2551 25.2

4 id 798 7.88

5 Warengruppe 798 7.88

6 Umsatz 798 7.88

7 Feiertag 394 3.89

8 Bewoelkung 84 0.829

9 Temperatur 29 0.286

10 Windgeschwindigkeit 29 0.286

11 Datum 0 0

12 Wochenende 0 0