

Package: MSBStatsData (via r-universe)

October 11, 2026

Title Data Sets for Courses at the Münster School of Business

Version 0.1.0

Description Provides sample data sets that are used in statistics and data science courses at the Münster School of Business. The datasets refer to different business topics but also other domains, e.g. sports, traffic, etc.

License GPL (>= 3)

Encoding UTF-8

LazyData true

LazyDataCompression xz

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.3

Depends R (>= 3.5)

Imports tibble

Suggests pkgdown

Repository <https://mchlbckr.r-universe.dev>

Date/Publication 2026-04-13 14:45:42 UTC

RemoteUrl <https://github.com/mchlbckr/msbstatsdata>

RemoteRef HEAD

RemoteSha 97a5cfc0050cc049012f0f3a956162d04ad215b6

Contents

berlin_marathon	3
beverage_revenues	4
bvb_rankings	4
car_occupation	5
car_occupation_ind	5
checkout_times	6
cinema_group_size_pmf	6

cinema_visitors	7
cinema_visitors_ind	7
cold_rents	8
company_financials	8
computer_sales_country	9
credit_ratings_two_reviewers	9
de_elections	10
de_energy_prices	10
decathlon	11
decathlon_3disc_covariances	11
decathlon_3disc_summary	12
env_survey_eu10	12
exam_study_time	13
florida_sentencing	13
florida_sentencing_ind	14
gas_prices	14
hr_data	15
library_overdue_model	15
library_overdue_sample	16
marketing_expenses	16
ms_bike_hourly_2025	17
ms_bike_hourly_2025_city	17
ms_bike_sites	18
nrw_population	18
order_quantities	19
populations_i_ii	19
product_demand_testphases	20
property_prices_distance	20
rent_living_area	21
research_ads	21
salary_trends	22
security_prices_ab	22
shipping_errors	23
sorting_times	23
sport_occupation	24
sport_occupation_ind	24
tv_gender	25
tv_gender_ind	25
union_satisfaction	26
union_satisfaction_ind	26
utility_vehicle_sales_weekly	27
vehicle_range_coefficients	27
website_dwell	28
website_dwell_ind	28
workshop_cars_per_hour	29
xy_quadratic	29

berlin_marathon	<i>Berlin Marathon results from 1999 to 2019</i>
-----------------	--

Description

Official finisher results for the Berlin Marathon from 1999 to 2019, including cumulative split times at 5 km intervals and half marathon.

Usage

```
berlin_marathon
```

Format

A tibble with 678,711 rows and 18 variables:

year Race year.

race Race name.

place_overall Overall finishing place in the published result list.

first_name Runner first name.

last_name Runner last name.

nationality Three-letter nationality code.

club Club or country/team label as published in the source data.

gender Runner gender ("M" or "W" in the source data).

time_full Official marathon finishing time.

split_5k Cumulative time at 5 km.

split_10k Cumulative time at 10 km.

split_15k Cumulative time at 15 km.

split_20k Cumulative time at 20 km.

time_half Cumulative half-marathon time.

split_25k Cumulative time at 25 km.

split_30k Cumulative time at 30 km.

split_35k Cumulative time at 35 km.

split_40k Cumulative time at 40 km.

Source

Data files from Andrew Miller's marathon-results repository: <https://github.com/AndrewMillerOnline/marathon-results>.

Berlin Marathon files used for this dataset: <https://github.com/AndrewMillerOnline/marathon-results/tree/main/Berlin>.

beverage_revenues	<i>Beverage revenues</i>
-------------------	--------------------------

Description

Annual revenues of two beverage manufacturers in thousand EUR.

Usage

beverage_revenues

Format

A tibble with 4 rows and 3 variables:

year Calendar year.

spritz Revenue of Spritz in thousand EUR.

prickli Revenue of Prickli in thousand EUR.

bvb_rankings	<i>Borussia Dortmund final rankings (1988-2022)</i>
--------------	---

Description

The dataset contains Borussia Dortmund's final Bundesliga ranking for each season from 1988 to 2022.

Usage

bvb_rankings

Format

A tibble with 35 rows and 2 variables:

year Calendar year of the season endpoint.

ranking Final Bundesliga ranking (1 = best rank).

car_occupation	<i>Car brand by occupation (aggregated)</i>
----------------	---

Description

The dataset contains counts of preferred car brands by occupation.

Usage

car_occupation

Format

A tibble with 4 rows and 6 variables:

occupation Occupation category.

audi Count of Audi drivers.

bmw Count of BMW drivers.

opel Count of Opel drivers.

vw Count of VW drivers.

total Row total.

car_occupation_ind	<i>Car brand by occupation (individual-level)</i>
--------------------	---

Description

The dataset contains one row per surveyed person with occupation and car brand.

Usage

car_occupation_ind

Format

A tibble with 500 rows and 2 variables:

occupation Occupation category.

brand Car brand category.

checkout_times	<i>Checkout service times</i>
----------------	-------------------------------

Description

The dataset contains service times (in seconds) of 50 consecutive customers at a supermarket checkout.

Usage

checkout_times

Format

A tibble with 50 rows and 1 variable:

service_time_seconds Observed service time in seconds.

cinema_group_size_pmf	<i>Cinema group size probability function</i>
-----------------------	---

Description

Probability function for group size in cinema visits.

Usage

cinema_group_size_pmf

Format

A tibble with 5 rows and 2 variables:

group_size Number of persons per group.

probability Probability mass.

cinema_visitors	<i>Cinema visitors over 100 days</i>
-----------------	--------------------------------------

Description

The dataset contains an absolute frequency distribution of cinema visitors per day recorded over 100 days.

Usage

```
cinema_visitors
```

Format

A tibble with 11 rows and 2 variables:

viewers Number of visitors counted on a day.

days Number of days with the respective visitor count.

cinema_visitors_ind	<i>Cinema visitors over 100 days (individual-level)</i>
---------------------	---

Description

Expanded version of [cinema_visitors](#) with one row per observed day.

Usage

```
cinema_visitors_ind
```

Format

A tibble with 100 rows and 1 variable:

viewers Number of visitors counted on the day.

cold_rents	<i>Monthly cold rents</i>
------------	---------------------------

Description

Ten observed monthly cold rents in EUR from one neighborhood.

Usage

```
cold_rents
```

Format

A tibble with 10 rows and 1 variable:

monthly_rent_eur Monthly cold rent in EUR.

company_financials	<i>Company profile sample</i>
--------------------	-------------------------------

Description

Sample of four companies used in introductory tasks on variables and scales.

Usage

```
company_financials
```

Format

A tibble with 4 rows and 5 variables:

company Company name.

employees Number of employees.

annual_revenue_mio_eur Annual revenue in million EUR.

equity_share_pct Equity share in percent.

credit_rating Rating category.

`computer_sales_country`*Monthly computer sales in countries A and B*

Description

Two discrete probability distributions for monthly sales volumes.

Usage`computer_sales_country`**Format**

A tibble with 7 rows and 3 variables:

units Monthly number of sold units.

probability_country_a Probability in country A.

probability_country_b Probability in country B.

`credit_ratings_two_reviewers`*Credit ratings from two reviewers*

Description

Two independent reviewers rated seven companies on a 1-10 scale.

Usage`credit_ratings_two_reviewers`**Format**

A tibble with 7 rows and 3 variables:

company_id Company identifier.

reviewer_a Score from reviewer A.

reviewer_b Score from reviewer B.

de_elections	<i>Bundestag election results by federal state</i>
--------------	--

Description

Party-level second-vote results by German federal state for the 2021 Bundestag election result table (official publication format).

Usage

```
de_elections
```

Format

A tibble with 377 rows and 6 variables:

election_year Election year.
state_code Two-digit federal state code.
state Federal state name.
party Party label in the official result file.
votes Number of second votes.
vote_share Second-vote share in percent.

de_energy_prices	<i>Monthly consumer energy price indices in Germany</i>
------------------	---

Description

Monthly Harmonised Index of Consumer Prices (HICP) sub-indices for selected energy categories in Germany (index base 2015 = 100).

Usage

```
de_energy_prices
```

Format

A tibble with 480 rows and 4 variables:

date Month of observation.
series Energy category series label.
coicop COICOP classification code.
price_index_2015_100 Price index value (2015 = 100).

decathlon	<i>Data of competition results of decathlon for men</i>
-----------	---

Description

The decathlon for men is a combined event in athletics consisting of 10 track and field events: 100 metres, 400 metres, 1500 metres, 110 metre hurdles, long jump, high jump, pole vault, discus throw, javelin throw, and shotput.

Usage

```
decathlon
```

Format

A tibble with 7,958 rows and 10 variables:

race100m dbl result of 100m race in seconds

longjump dbl result of long jump in meters

shotput dbl result of shot put in meters

highjump dbl result of high jump in meters

race400m dbl result of 400m race in seconds

race110mhurdles dbl result of 110m hurdles race in seconds

discus dbl result of discus throw in meters

polevault dbl result of pole vault in meters

javelinthrow dbl result of javelin throw in meters

race1500m dbl result of 1500m race in seconds

decathlon_3disc_covariances	<i>Decathlon covariances for three discipline pairs</i>
-----------------------------	---

Description

Covariances between the same three disciplines used in decathlon_3disc_summary.

Usage

```
decathlon_3disc_covariances
```

Format

A tibble with 3 rows and 2 variables:

pair Discipline pair label.

covariance Covariance for the pair.

 decathlon_3disc_summary

Decathlon summary for three disciplines

Description

Mean and variance for 100m, long jump, and shot put from a large decathlon sample.

Usage

```
decathlon_3disc_summary
```

Format

A tibble with 3 rows and 3 variables:

discipline Discipline name.

mean Arithmetic mean of results.

variance Variance of results.

 env_survey_eu10

Environmental survey sample (EU, n = 10)

Description

The dataset contains responses of 10 sampled persons from an EU-wide survey on environmental protection.

Usage

```
env_survey_eu10
```

Format

A tibble with 10 rows and 4 variables:

age Age of respondent.

gender Gender of respondent.

income_eur Monthly net income in EUR.

environmental_protection Attitude towards environmental protection.

exam_study_time	<i>Exam points and study time</i>
-----------------	-----------------------------------

Description

Study days and exam points of five participants.

Usage

```
exam_study_time
```

Format

A tibble with 5 rows and 3 variables:

study_days Study time in days.

points Achieved points.

max_points Maximum achievable points.

florida_sentencing	<i>Florida murder sentencing by defendant skin color (aggregated)</i>
--------------------	---

Description

The dataset contains a contingency table on sentencing outcomes by defendant skin color.

Usage

```
florida_sentencing
```

Format

A tibble with 3 rows and 4 variables:

defendant_skin_color Defendant skin color and total row.

death_penalty Count of death penalty sentences.

other_sentence Count of other sentences.

total Row total.

`florida_sentencing_ind`*Florida murder sentencing by defendant skin color (individual-level)*

Description

The dataset contains one row per case with defendant skin color and sentence category.

Usage`florida_sentencing_ind`**Format**

A tibble with 326 rows and 2 variables:

defendant_skin_color Defendant skin color.

sentence Sentence category.

`gas_prices`*Motorway gas prices*

Description

Ten observed prices for regular gasoline (EUR per liter) from a motorway section.

Usage`gas_prices`**Format**

A tibble with 10 rows and 1 variable:

price_eur_per_liter Observed price in EUR per liter.

hr_data*HR data*

Description

The dataset contains human resources data of a footwear company. Each entity of the dataset represents one employee with a total of six attributes.

Usage

hr_data

Format

A tibble with 1,200 rows and 6 variables:

position [fct] Position of the employee in the company.

working_hours [dbl] Contracted working hours of the employee per week.

salary [dbl] Monthly salary of the employee.

hourly_wage [dbl] Hourly compensation of the employee.

department [fct] Department in which the employee is employed.

sick_days [dbl] Sick days of the employee in the period observed.

library_overdue_model *Library overdue model*

Description

Symbolic probability model for overdue weeks.

Usage

library_overdue_model

Format

A tibble with 5 rows and 2 variables:

weeks_overdue Overdue weeks.

probability_expression Probability expression as text.

library_overdue_sample

Library overdue sample

Description

Frequency sample of overdue weeks used for method-of-moments exercises.

Usage

```
library_overdue_sample
```

Format

A tibble with 5 rows and 2 variables:

weeks_overdue Overdue weeks.

count Observed frequency.

marketing_expenses

Data of Marketing expenses

Description

The dataset contains product-, marketing- and sales data of 235 shoes of a footwear company. Each entity represents one shoe, listed with its product data, marketing data and sales data in a total of 14 variables.

Usage

```
marketing_expenses
```

Format

A tibble with 235 rows and 14 variables:

marketing_expenses [dbl] Expenses for marketing activities for the shoe.

customers_reached [dbl] Estimated number of customers reached by the footwear marketing activity of the shoe.

negative_reactions [dbl] Number of negative reactions to the marketing activities of the shoe.

price [dbl] Retail price of the shoe.

price_segment [fct] Price segment of the shoe.

number_of_sizes [dbl] Number of sizes in which the shoe is available.

target_customer [fct] Gender the shoe is intended for.

rating_testers [dbl] Average product rating of the test customers for the shoe.

rating_customers [dbl] Average product rating of the real customers for the shoe.

color_most_sold [fct] Color in which the shoe is sold the most often.

return_rate [dbl] Rate how often the shoe is returned by the customer.

sales_volume [dbl] Number of sales for the shoe.

rank_rating_customers [dbl] Attribute 'rating_customers', divided into ranks for the calculation of the correlation coefficient according to Spearman.

rank_price_segment [dbl] Attribute 'price_segment', divided into ranks for the calculation of the correlation coefficient according to Spearman.

ms_bike_hourly_2025 *Hourly bike counts in Muenster by station (2025)*

Description

Hourly totals derived from 15-minute open data published by the City of Muenster for available bike counting stations in 2025.

Usage

ms_bike_hourly_2025

Format

A tibble with 210,240 rows and 4 variables:

station_id Station identifier.

station_name Readable station name.

datetime_hour Hour timestamp (local time, Europe/Berlin).

bikes_total Total counted bikes in the hour.

ms_bike_hourly_2025_city *Hourly citywide bike counts in Muenster (2025)*

Description

Citywide hourly totals aggregated across available bike counting stations in Muenster for 2025.

Usage

ms_bike_hourly_2025_city

Format

A tibble with 8,760 rows and 2 variables:

datetime_hour Hour timestamp (local time, Europe/Berlin).

bikes_total Total counted bikes across all available stations.

ms_bike_sites	<i>Bike counting sites in Muenster</i>
---------------	--

Description

Metadata for selected public bike counting stations in Muenster.

Usage

ms_bike_sites

Format

A tibble with 24 rows and 2 variables:

station_id Station identifier used in the source repository.

station_name Readable station name.

nrw_population	<i>Population by NRW districts</i>
----------------	------------------------------------

Description

Annual population counts for NUTS-3 districts in North Rhine-Westphalia (NRW), Germany.

Usage

nrw_population

Format

A tibble with 318 rows and 5 variables:

year Reference year.

nuts3_code NUTS-3 district code.

district_name District name.

district_type District type (urban or rural district).

population Total resident population.

order_quantities	<i>Order quantities</i>
------------------	-------------------------

Description

Order quantities per transaction from a small ordering sample.

Usage

order_quantities

Format

A tibble with 20 rows and 1 variable:

order_quantity Ordered units per order.

populations_i_ii	<i>Two populations I and II</i>
------------------	---------------------------------

Description

Two small populations used for comparing dispersion measures.

Usage

populations_i_ii

Format

A tibble with 4 rows and 2 variables:

population_i Values of population I.

population_ii Values of population II.

product_demand_testphases

Product demand test phases

Description

Demand measured across test phases with varying selling price and product quality.

Usage

product_demand_testphases

Format

A tibble with 6 rows and 3 variables:

selling_price_eur Selling price in EUR.

product_quality Product quality score.

demand Observed demand.

property_prices_distance

Property prices and station distance

Description

Property prices per square meter and distance to next suburban rail station.

Usage

property_prices_distance

Format

A tibble with 5 rows and 2 variables:

distance_km Distance to station in km.

price_eur_per_m2 Price in EUR per square meter.

rent_living_area	<i>Rent and living area</i>
------------------	-----------------------------

Description

Monthly cold rent and living area for five apartments.

Usage

rent_living_area

Format

A tibble with 5 rows and 2 variables:

living_area_m2 Living area in square meters.

cold_rent_eur Monthly cold rent in EUR.

research_ads	<i>Monthly research and advertising spending</i>
--------------	--

Description

The dataset contains monthly spending of a company on research and advertising (in thousand EUR).

Usage

research_ads

Format

A tibble with 12 rows and 3 variables:

month Month abbreviation.

research Research spending in thousand EUR.

advertising Advertising spending in thousand EUR.

salary_trends	<i>Salary trend data</i>
---------------	--------------------------

Description

Dataset containing the average salary of a footwear company's employees over 10 years.

Usage

salary_trends

Format

A tibble with 10 rows and 2 variables:

year [dbl] Year of record.

avg_salary [dbl] Average salary in the corresponding year.

security_prices_ab	<i>Annual closing prices of two securities</i>
--------------------	--

Description

Five annual closing prices for security A and security B.

Usage

security_prices_ab

Format

A tibble with 5 rows and 3 variables:

year Time index (1 to 5).

paper_a Closing price of security A in EUR.

paper_b Closing price of security B in EUR (year 5 unknown = NA).

shipping_errors	<i>Shipping error data</i>
-----------------	----------------------------

Description

Dataset recording errors in the shipment and the type of delivery requested by the customer in the process.

Usage

shipping_errors

Format

A tibble with 576 rows and 2 variables:

error [fct] Indicator whether or not an error appeared during shipment.

shipping [fct] Shipment method requested by the customer.

sorting_times	<i>Sorting times</i>
---------------	----------------------

Description

The dataset contains throughput times of a machine recorded in order to compare them to the manufacturer's specifications.

Usage

sorting_times

Format

A tibble with 60 rows and 1 variable:

sorting_time [dbl] Recorded time (in seconds) required by the machine for sorting individual parts.

sport_occupation	<i>Sport activity by occupation (aggregated)</i>
------------------	--

Description

The dataset contains counts of sport activity frequencies by occupation.

Usage

```
sport_occupation
```

Format

A tibble with 5 rows and 4 variables:

occupation Occupation category.

never Count with no sport activity.

occasionally Count with occasional sport activity.

regularly Count with regular sport activity.

sport_occupation_ind	<i>Sport activity by occupation (individual-level)</i>
----------------------	--

Description

The dataset contains one row per surveyed person with occupation and sport activity frequency.

Usage

```
sport_occupation_ind
```

Format

A tibble with 1,000 rows and 2 variables:

occupation Occupation category.

sport_activity Sport activity frequency category.

tv_gender	<i>TV viewing time by gender (aggregated)</i>
-----------	---

Description

The dataset contains grouped TV viewing counts by gender.

Usage

tv_gender

Format

A tibble with 3 rows and 4 variables:

tv_time_group Grouped weekly TV viewing time and total row.

men Count of men.

women Count of women.

total Row total.

tv_gender_ind	<i>TV viewing time by gender (individual-level)</i>
---------------	---

Description

The dataset contains one row per person with gender and grouped weekly TV viewing time.

Usage

tv_gender_ind

Format

A tibble with 320 rows and 2 variables:

gender Gender of the surveyed person.

tv_time_group Grouped weekly TV viewing time.

union_satisfaction	<i>Satisfaction with a union agreement</i>
--------------------	--

Description

The dataset contains a contingency table on member satisfaction with a new agreement for two unions (plus totals).

Usage

```
union_satisfaction
```

Format

A tibble with 3 rows and 6 variables:

union Union identifier and total row.

very_satisfied Count of very satisfied respondents.

satisfied Count of satisfied respondents.

dissatisfied Count of dissatisfied respondents.

very_dissatisfied Count of very dissatisfied respondents.

total Row total.

union_satisfaction_ind	<i>Satisfaction with a union agreement (individual-level)</i>
------------------------	---

Description

The dataset contains the same information as [union_satisfaction](#), expanded to individual-level observations with one row per surveyed person.

Usage

```
union_satisfaction_ind
```

Format

A tibble with 1,000 rows and 2 variables:

union Union identifier.

satisfaction Satisfaction category for the agreement.

utility_vehicle_sales_weekly
<i>Weekly utility vehicle sales</i>

Description

Probability function for weekly units sold of utility vehicles.

Usage

utility_vehicle_sales_weekly

Format

A tibble with 6 rows and 2 variables:

vehicles_sold Number of vehicles sold per week.

probability Probability mass.

vehicle_range_coefficients
<i>Vehicle range model coefficients</i>

Description

Estimated coefficients from a multiple linear regression for vehicle range.

Usage

vehicle_range_coefficients

Format

A tibble with 4 rows and 2 variables:

regressor Regressor name.

coefficient Estimated coefficient.

website_dwell	<i>Grouped website dwell times</i>
---------------	------------------------------------

Description

The dataset contains grouped dwell-time frequencies from a sample of 100 website users.

Usage

```
website_dwell
```

Format

A tibble with 5 rows and 2 variables:

dwell_time_interval_min Interval of dwell time in minutes.

persons Number of persons in the interval.

website_dwell_ind	<i>Grouped website dwell times (individual-level)</i>
-------------------	---

Description

Expanded version of [website_dwell](#) with one row per sampled person.

Usage

```
website_dwell_ind
```

Format

A tibble with 100 rows and 1 variable:

dwell_time_interval_min Interval of dwell time in minutes for the person.

workshop_cars_per_hour	<i>Cars serviced per hour</i>
------------------------	-------------------------------

Description

Probability function for hourly number of cars serviced in a repair shop.

Usage

workshop_cars_per_hour

Format

A tibble with 3 rows and 2 variables:

cars_per_hour Number of cars serviced per hour.

probability Probability mass.

xy_quadratic	<i>Quadratic x-y observations</i>
--------------	-----------------------------------

Description

Seven observations with a symmetric quadratic relationship between x and y.

Usage

xy_quadratic

Format

A tibble with 7 rows and 3 variables:

i Observation index.

x Value of X.

y Value of Y.

Index

* datasets

berlin_marathon, [3](#)
beverage_revenues, [4](#)
bvb_rankings, [4](#)
car_occupation, [5](#)
car_occupation_ind, [5](#)
checkout_times, [6](#)
cinema_group_size_pmf, [6](#)
cinema_visitors, [7](#)
cinema_visitors_ind, [7](#)
cold_rents, [8](#)
company_financials, [8](#)
computer_sales_country, [9](#)
credit_ratings_two_reviewers, [9](#)
de_elections, [10](#)
de_energy_prices, [10](#)
decathlon, [11](#)
decathlon_3disc_covariances, [11](#)
decathlon_3disc_summary, [12](#)
env_survey_eu10, [12](#)
exam_study_time, [13](#)
florida_sentencing, [13](#)
florida_sentencing_ind, [14](#)
gas_prices, [14](#)
hr_data, [15](#)
library_overdue_model, [15](#)
library_overdue_sample, [16](#)
marketing_expenses, [16](#)
ms_bike_hourly_2025, [17](#)
ms_bike_hourly_2025_city, [17](#)
ms_bike_sites, [18](#)
nrw_population, [18](#)
order_quantities, [19](#)
populations_i_ii, [19](#)
product_demand_testphases, [20](#)
property_prices_distance, [20](#)
rent_living_area, [21](#)
research_ads, [21](#)
salary_trends, [22](#)
security_prices_ab, [22](#)
shipping_errors, [23](#)
sorting_times, [23](#)
sport_occupation, [24](#)
sport_occupation_ind, [24](#)
tv_gender, [25](#)
tv_gender_ind, [25](#)
union_satisfaction, [26](#)
union_satisfaction_ind, [26](#)
utility_vehicle_sales_weekly, [27](#)
vehicle_range_coefficients, [27](#)
website_dwell, [28](#)
website_dwell_ind, [28](#)
workshop_cars_per_hour, [29](#)
xy_quadratic, [29](#)

berlin_marathon, [3](#)
beverage_revenues, [4](#)
bvb_rankings, [4](#)

car_occupation, [5](#)
car_occupation_ind, [5](#)
checkout_times, [6](#)
cinema_group_size_pmf, [6](#)
cinema_visitors, [7](#), [7](#)
cinema_visitors_ind, [7](#)
cold_rents, [8](#)
company_financials, [8](#)
computer_sales_country, [9](#)
credit_ratings_two_reviewers, [9](#)

de_elections, [10](#)
de_energy_prices, [10](#)
decathlon, [11](#)
decathlon_3disc_covariances, [11](#)
decathlon_3disc_summary, [12](#)

env_survey_eu10, [12](#)
exam_study_time, [13](#)

florida_sentencing, [13](#)

florida_sentencing_ind, 14

gas_prices, 14

hr_data, 15

library_overdue_model, 15

library_overdue_sample, 16

marketing_expenses, 16

ms_bike_hourly_2025, 17

ms_bike_hourly_2025_city, 17

ms_bike_sites, 18

nrw_population, 18

order_quantities, 19

populations_i_ii, 19

product_demand_testphases, 20

property_prices_distance, 20

rent_living_area, 21

research_ads, 21

salary_trends, 22

security_prices_ab, 22

shipping_errors, 23

sorting_times, 23

sport_occupation, 24

sport_occupation_ind, 24

tv_gender, 25

tv_gender_ind, 25

union_satisfaction, 26, 26

union_satisfaction_ind, 26

utility_vehicle_sales_weekly, 27

vehicle_range_coefficients, 27

website_dwell, 28, 28

website_dwell_ind, 28

workshop_cars_per_hour, 29

xy_quadratic, 29