

# Deterministic Model 01: Příprava dat (Data Preparation)

```
In [1]: # Instalace potřebných knihoven
        %%pip install pandas
        %%pip install numpy
```

```
In [2]: # Import potřebných knihoven
        import pandas as pd
        import numpy as np
```

## Načtení reálných dat a dat digitálního kladečského plánu

```
In [3]: # Soubor je načten a přiřazen do proměnné ,df"
        other_path = '../..data/01_DataScience/final_timelaps.csv'
        df = pd.read_csv(other_path, header=0)
```

```
In [4]: # Zobrazení prvních 5 řádků datasetu
        print('Prvních 5 řádků datového rámce')
        df.head(5)
```

Prvních 5 řádků datového rámce

```
Out[4]:
```

	id	x	y	z	time	delay	type_delay	total_time
0	150	1315	220	1000	29	0	0	29
1	75	220	1190	500	33	0	0	33
2	239	220	940	2000	35	6	3	41
3	199	1315	220	1500	36	0	0	36
4	51	3690	220	250	50	0	0	50

```
In [5]: # Zobrazení prvních 5 řádků datasetu
        print('Prvních 5 řádků datového rámce')
        df.head(5)
```

Prvních 5 řádků datového rámce

```
Out[5]:
```

	id	x	y	z	time	delay	type_delay	total_time
0	150	1315	220	1000	29	0	0	29
1	75	220	1190	500	33	0	0	33
2	239	220	940	2000	35	6	3	41
3	199	1315	220	1500	36	0	0	36
4	51	3690	220	250	50	0	0	50

```
In [6]: # Kontrola
        df.dtypes
```

```
Out[6]: id          int64
        x          int64
        y          int64
        z          int64
        time       int64
        delay      int64
        type_delay  int64
        total_time int64
        dtype: object
```

```
In [7]: df.describe()
```

Out[7]:		id	x	y	z	time	delay	type_delay	total_ti
	count	161.000000	161.000000	161.000000	161.000000	161.000000	161.000000	161.000000	161.000000
	mean	139.049689	1314.428571	1091.807453	993.788820	36.285714	5.708075	0.254658	41.993750
	std	79.374177	1288.402896	1361.839636	702.645843	6.771658	25.355384	0.800621	27.037500
	min	2.000000	95.000000	95.000000	0.000000	24.000000	0.000000	0.000000	24.000000
	25%	71.000000	220.000000	220.000000	500.000000	32.000000	0.000000	0.000000	32.000000
	50%	141.000000	690.000000	220.000000	1000.000000	35.000000	0.000000	0.000000	36.000000
	75%	203.000000	2440.000000	1565.000000	1500.000000	39.000000	0.000000	0.000000	42.000000
	max	277.000000	4002.000000	4690.000000	2250.000000	56.000000	200.000000	4.000000	250.000000

## Export datové sady do formátu CSV

```
In [8]: df.to_csv('../data/02_DetermModel/real_data.csv', index=False)
```

## Autor / Organizace / Datum

Vjačeslav Usmanov, ČVUT v Praze, Fakulta stavební

Přehled změn

Datum (YYYY-MM-DD)	Verze	Autor změny	Popis změny
2026-01-20	1.1	Vjačeslav Usmanov	added DM_01_Data_Preparing.ipynb
2026-02-12	1.2	Vjačeslav Usmanov	changed DM_01_Data_Preparing.ipynb