✅ Your analysis script for jahresformation.csv executed successfully and revealed a large, rich dataset with **383,354 rows and 30 columns**.

**🧠 Key insights:**

**✅ Core fields for linking & route construction**

| **Column** | **Description (inferred)** |
| --- | --- |
| Train | Likely a unique identifier per train run or service instance |
| Train type | Service category (e.g., R, IC, IR) |
| From station / To station | Start and end stations |
| Block designation | Key to link with rollmaterial.csv for vehicle types |
| Start of timetable period | Beginning of the schedule validity |
| TP-daily runs (Column R) | Descriptive tags of operation days (e.g., Mo–So) |
| Bitmap | Bit-encoded run-day pattern (can be decoded into calendar) |

**⚠️ Missing values**

* Suffix, Traction: Almost entirely null (may be dropped or inferred)
* From.1 and By.1: Only available for ~64% of entries — double-check use case.
* Bitmap is null in **223 rows**, may affect decoding schedules.
* Most other fields are complete.

**💡 Next useful actions:**

1. **Decode Bitmap** to understand actual operation days for each train.
2. **Map Block designation to rollmaterial.csv** using your type conversion file.
3. Use From station, To station, Train, and Start of timetable period to build GTFS-equivalent route representations for SUMO.

Would you like me to help you:

* Decode the Bitmap into run-day flags?
* Create a clean version of this dataset with only useful, de-duplicated fields?
* Or generate a mapping of Block designation to actual rolling stock entries?

Let me know which route you want to explore next.