**Key Insights from the Deals Data**

**1. Only Open Deals Have a Deal Score**

* deal\_score is only available for deals that are still open.
* Closed deals (both Won and Lost) don’t have a score.
* ➤ **Action**: Use is\_closed\_won as the true outcome variable for win/loss analysis.

**2. Closed Won Deals Exist, but Deal Score Is Missing**

* You found **158 deals marked as Closed Won** — useful for success analysis.
* But none of them have a deal\_score.
* ➤ **Implication**: To use deal\_score as a predictor, focus on open deals and future outcomes.

**3. Smaller Deals Win More Often**

* Deals < $10K win more frequently.
* Enterprise deals (>$100K) are harder to close but more valuable.
* ➤ **Implication**: Balance **volume vs. value** in sales prioritization.

**4. Deal Source Matters**

* Some sources (e.g., **Referral Partner**, **Event**) perform better than others.
* ➤ **Implication**: Double down on **high-converting channels**.

**5. Sales Cycle & Pipeline Time Vary by Outcome**

* Won deals typically have **shorter pipeline durations**.
* You created total\_pipeline\_time\_sec to track this.
* ➤ **Insight**: **Longer pipelines** may indicate deals at risk.

**6. Feature Engineering Is Strong**

Feature engineered several useful features:

* Time in stages (secs\_bant, secs\_opportunity, etc.)
* Deal size categories
* Sales cycle categories
* Total pipeline time
* Clean is\_closed\_won and is\_won fields

➤ **These will be valuable inputs for future modeling or dashboards.**

* **Stage Time (secs\_bant, secs\_opportunity, etc.):** Converted time strings to seconds to measure how long deals spend in each pipeline stage.
* **Deal Size Category:** Grouped deals into Small, Medium, Large, and Enterprise to compare performance by size.
* **Sales Cycle Category:** Labeled deals based on how many days they took to close (e.g., Short, Medium, Long).
* **Total Pipeline Time:** Created a single metric that adds up time across all stages.
* **Cleaned Success Flags:** Fixed is\_closed\_won and created is\_won for clear win/loss tracking.

|  |  |
| --- | --- |
| Feature | Why It's Useful |
| weighted\_amount | Are bigger/smaller deals more successful? |
| deal\_size\_category | Are small, medium, or enterprise deals more likely to close? |
| deal\_type | Do renewals close more than new deals? |
| deal\_source\_attribution\_2 | Which sources have the best win rates? |
| total\_pipeline\_time\_sec | Are faster-moving deals more likely to close? |
| contract\_term\_months | Any pattern with length of contracts? |
| create\_year\_month | Are there time-based trends in success? |