

## Analyzing Customer Payment Behaviour to Improve Collections Efficiency and Working Capital (SQL Case Study)

### Business question :

- The objective was to evaluate how customer payment behaviour impacts collection efficiency, Days Sales Outstanding (DSO), and cash-flow predictability, and to identify segments contributing to delayed realization.

### Use Cases :

- Identify whether delayed collections are customer-driven or process-driven. Detect industries or portfolios contributing disproportionately to overdue receivables.
- Support prioritization of high-risk accounts for collection follow-up. Improve short-term cash forecasting through behavioural analysis

### Dataset Used :

- The dataset represents invoice-level Accounts Receivable activity over a 6-month period, simulating a typical enterprise collections ledger.
- Some the fields used are as below
  - o Customer\_name
  - o Invoice\_number
  - o Invoice\_date
  - o Due\_date
  - o Payment\_date
  - o Industry
  - o Age\_bucket

### Assumption :

- Data used in the dataset is anonymous and does not indicate any actual customer details

**Analytics :** Four derived key metrics were calculated to identify company's working capital management.

- **Weighted DSO by Industry :** For accurate calculation of DSO eliminating variances in NET terms, Invoice amount.
  - o **SQL Query used :**
  - o `SELECT`
  - o `industry,`
  - o `SUM(invoice_amount * delay_days) / SUM(invoice_amount) AS weighted_dso`
  - o `FROM (`
  - o `SELECT`
  - o `industry,`
  - o `invoice_amount,`
  - o `DATEDIFF(payment_date, invoice_date) AS delay_days`
  - o `FROM oracle_ar`
  - o `WHERE payment_date IS NOT NULL`
  - o `) t`
  - o `GROUP BY industry;`

- **Aging Risk Distribution** : To identify the cash in age buckets, and understand the risk associated with it.

- o **SQL Query used :**
- o *SELECT*
- o *age\_bucket,*
- o *SUM(invoice\_amount) AS total\_ar,*
- o *ROUND(*
- o *SUM(invoice\_amount) / SUM(SUM(invoice\_amount)) OVER () \* 100,*
- o *2*
- o *) AS open\_ar\_pct*
- o *FROM oracle\_ar*
- o *WHERE payment\_date IS NULL*
- o *GROUP BY age\_bucket;*

- **Pareto Analysis (Customer concentration)** : Helps to prioritize Collection efforts and manage resources.

- o **SQL Query Used :**
- o *SELECT*
- o *customer\_name,*
- o *total\_ar,*
- o *RANK() OVER (ORDER BY total\_ar DESC) AS risk\_rank,*
- o *SUM(total\_ar) OVER (ORDER BY total\_ar DESC) AS cumu\_exp,*
- o *ROUND(*
- o *SUM(total\_ar) OVER (ORDER BY total\_ar DESC) / SUM(total\_ar) OVER () \* 100, 2*
- o *) AS cumu\_pct*
- o *FROM (*
- o *SELECT*
- o *customer\_name,*
- o *SUM(invoice\_amount) AS total\_ar*
- o *FROM oracle\_ar*
- o *WHERE payment\_date IS NULL*
- o *GROUP BY customer\_name*
- o *) t;*

- **Chronic Late-Payer Identification** : Used to identify structural payment issues rather than isolated incidents.

- o **SQL Query Used :**
- o *SELECT*
- o *customer\_name,*
- o *COUNT(\*) AS overdue\_invoices,*
- o *ROUND(*
- o *AVG(DATEDIFF(COALESCE(payment\_date, CURDATE()), due\_date)), 1*
- o *) AS avg\_delay\_days,*
- o *SUM(invoice\_amount) AS overdue\_exposure*
- o *FROM oracle\_ar*
- o *WHERE DATEDIFF(COALESCE(payment\_date, CURDATE()), due\_date) > 0*
- o *GROUP BY customer\_name;*

### Findings and Interpretation :

- **Weighted DSO** analysis showed that delays are not coming from small invoices but from higher-value ones, meaning a few large payments getting delayed is stretching the company's cash cycle.
- **Aging analysis** made it clear that a big share of receivables has already moved into later buckets (60+ / 90+), which suggests this is not just timing — collections are slowing down over time.
- **Customer concentration** check showed that AR is spread across many customers rather than stuck with a few, so the issue is more process-driven than customer-specific risk.
- **Behavioural analysis** highlighted that some customers consistently pay late, indicating a pattern rather than one-off delays.

### Business Implications :

- Collections teams should focus more on timely follow-ups for large-value invoices since they influence cash flow the most.
- There is a need to tighten internal processes like dispute closure, billing clarity, and reminder cycles to prevent invoices from aging further.
- Since exposure is not concentrated, improving the overall collections workflow will have more impact than chasing only a few customers.
- Customers showing repeated delays may need revised payment terms or closer monitoring to avoid future risk buildup.

### Conclusion :

- The analysis demonstrates how invoice-level data can be leveraged using SQL to evaluate collections efficiency and receivables risk.
- Combining DSO measurement, aging distribution, exposure concentration, and payment behaviour provides a comprehensive view of receivables performance.
- Such analysis supports more informed decision-making in working capital management and collections prioritization.
- Regular monitoring of these metrics can help organizations maintain healthier cash-flow cycles and identify emerging risks earlier.