PySpark — Spark SQL Only Problems

problem_001

```
-- Problem 001: 001 - Aggregations & GroupBy: Approx_count_distinct challenge
-- Category: Aggregations & GroupBy

-- Register temp view in PySpark:
-- transactions.createOrReplaceTempView('transactions_view');

SELECT *
FROM transactions_view
LIMIT 100;
```

problem_002

```
-- Problem 002: 002 - Pivot & Crosstab: Cube challenge
-- Category: Pivot & Crosstab

-- Register temp view in PySpark:
-- transactions.createOrReplaceTempView('transactions_view');

SELECT *
FROM transactions_view
LIMIT 100;
```

problem_003

```
-- Problem 003: 003 - Strings & Regex: Concat_ws challenge
-- Category: Strings & Regex

-- Register temp view in PySpark:
-- orders.createOrReplaceTempView('orders_view');

SELECT *
FROM orders_view
LIMIT 100;
```

problem_004

```
-- Problem 004: 004 - Aggregations & GroupBy: Groupby challenge
-- Category: Aggregations & GroupBy

-- Register temp view in PySpark:
-- logs.createOrReplaceTempView('logs_view');

SELECT *
FROM logs_view
LIMIT 100;
```

problem_006

```
-- Problem 006: 006 - Aggregations & GroupBy: Groupby challenge
-- Category: Aggregations & GroupBy

-- Register temp view in PySpark:
-- sessions.createOrReplaceTempView('sessions_view');

SELECT *
FROM sessions_view
LIMIT 100;
```

problem_007

```
-- Problem 007: 007 - Dates & Timestamps: Date_trunc challenge
-- Category: Dates & Timestamps

-- Register temp view in PySpark:
-- transactions.createOrReplaceTempView('transactions_view');

SELECT *
FROM transactions_view
LIMIT 100;
```

problem_010

```
-- Problem 010: 010 - Pivot & Crosstab: Cube challenge
-- Category: Pivot & Crosstab

-- Register temp view in PySpark:
-- logs.createOrReplaceTempView('logs_view');

SELECT *
FROM logs_view
LIMIT 100;
```

problem_011

```
-- Problem 011: 011 - Joins: Semi challenge
-- Category: Joins

-- Register temp view in PySpark:
-- events.createOrReplaceTempView('events_view');

SELECT *
FROM events_view
LIMIT 100;
```

problem_012

```
-- Problem 012: 012 - Aggregations & GroupBy: Approx_count_distinct challenge
-- Category: Aggregations & GroupBy

-- Register temp view in PySpark:
-- logs.createOrReplaceTempView('logs_view');

SELECT *
FROM logs_view
LIMIT 100;
```

problem_013

```
-- Problem 013: 013 - Window Functions: Dense_rank challenge
-- Category: Window Functions

-- Register temp view in PySpark:
-- orders.createOrReplaceTempView('orders_view');

SELECT *
FROM orders_view
LIMIT 100;
```

problem_015

```
-- Problem 015: 015 - Pivot & Crosstab: Cube challenge
-- Category: Pivot & Crosstab

-- Register temp view in PySpark:
-- orders.createOrReplaceTempView('orders_view');

SELECT *
FROM orders_view
LIMIT 100;
```

problem_016

```
-- Problem 016: 016 - Spark SQL: Functions in sql challenge
-- Category: Spark SQL

-- Register temp view in PySpark:
-- orders.createOrReplaceTempView('orders_view');

SELECT *
FROM orders_view
LIMIT 100;
```

problem_017

```
-- Problem 017: 017 - Pivot & Crosstab: Cube challenge
-- Category: Pivot & Crosstab

-- Register temp view in PySpark:
-- orders.createOrReplaceTempView('orders_view');

SELECT *
FROM orders_view
LIMIT 100;
```

problem_018

```
-- Problem 018: 018 - Window Functions: Lead challenge
-- Category: Window Functions

-- Register temp view in PySpark:
-- users.createOrReplaceTempView('users_view');

SELECT *
FROM users_view
LIMIT 100;
```

problem_019

```
-- Problem 019: 019 - Window Functions: Rank challenge
-- Category: Window Functions

-- Register temp view in PySpark:
-- transactions.createOrReplaceTempView('transactions_view');

SELECT *
FROM transactions_view
LIMIT 100;
```

problem_021

```
-- Problem 021: 021 - Dates & Timestamps: Date_trunc challenge
-- Category: Dates & Timestamps

-- Register temp view in PySpark:
-- clicks.createOrReplaceTempView('clicks_view');

SELECT *
FROM clicks_view
LIMIT 100;
```

problem_022

```
-- Problem 022: 022 - Pivot & Crosstab: Pivot challenge
-- Category: Pivot & Crosstab

-- Register temp view in PySpark:
-- users.createOrReplaceTempView('users_view');

SELECT *
FROM users_view
LIMIT 100;
```

problem_023

```
-- Problem 023: 023 - Dates & Timestamps: Date_trunc challenge
-- Category: Dates & Timestamps

-- Register temp view in PySpark:
-- events.createOrReplaceTempView('events_view');

SELECT *
FROM events_view
LIMIT 100;
```

problem_024

```
-- Problem 024: 024 - Pivot & Crosstab: Rollup challenge
-- Category: Pivot & Crosstab

-- Register temp view in PySpark:
-- logs.createOrReplaceTempView('logs_view');

SELECT *
FROM logs_view
LIMIT 100;
```

problem_026

```
-- Problem 026: 026 - Spark SQL: Sql queries challenge
-- Category: Spark SQL

-- Register temp view in PySpark:
-- clicks.createOrReplaceTempView('clicks_view');

SELECT *
FROM clicks_view
LIMIT 100;
```

problem_028

```
-- Problem 028: 028 - Window Functions: Rank challenge
-- Category: Window Functions

-- Register temp view in PySpark:
-- clicks.createOrReplaceTempView('clicks_view');

SELECT *
FROM clicks_view
LIMIT 100;
```

problem_029

```
-- Problem 029: 029 - Window Functions: Row_number challenge
-- Category: Window Functions

-- Register temp view in PySpark:
-- orders.createOrReplaceTempView('orders_view');

SELECT *
FROM orders_view
LIMIT 100;
```

problem_032

```
-- Problem 032: 032 - Spark SQL: Sql queries challenge
-- Category: Spark SQL

-- Register temp view in PySpark:
-- users.createOrReplaceTempView('users_view');

SELECT *
FROM users_view
LIMIT 100;
```

problem_033

```
-- Problem 033: 033 - Pivot & Crosstab: Pivot challenge
-- Category: Pivot & Crosstab

-- Register temp view in PySpark:
-- clicks.createOrReplaceTempView('clicks_view');

SELECT *
FROM clicks_view
LIMIT 100;
```

problem_035

```
-- Problem 035: 035 - Pivot & Crosstab: Cube challenge
-- Category: Pivot & Crosstab

-- Register temp view in PySpark:
-- logs.createOrReplaceTempView('logs_view');

SELECT *
FROM logs_view
LIMIT 100;
```

problem_037

```
-- Problem 037: 037 - Dates & Timestamps: From_unixtime challenge
-- Category: Dates & Timestamps

-- Register temp view in PySpark:
-- products.createOrReplaceTempView('products_view');

SELECT *
FROM products_view
LIMIT 100;
```

problem_038

```
-- Problem 038: 038 - Pivot & Crosstab: Rollup challenge
-- Category: Pivot & Crosstab

-- Register temp view in PySpark:
-- users.createOrReplaceTempView('users_view');

SELECT *
FROM users_view
LIMIT 100;
```

problem_043

```
-- Problem 043: 043 - Joins: Semi challenge
-- Category: Joins

-- Register temp view in PySpark:
-- logs.createOrReplaceTempView('logs_view');

SELECT *
FROM logs_view
LIMIT 100;
```

problem_051

```
-- Problem 051: 051 - Window Functions: Dense_rank challenge
-- Category: Window Functions

-- Register temp view in PySpark:
-- logs.createOrReplaceTempView('logs_view');

SELECT *
FROM logs_view
LIMIT 100;
```

problem_052

```
-- Problem 052: 052 - Aggregations & GroupBy: Agg challenge
-- Category: Aggregations & GroupBy

-- Register temp view in PySpark:
-- users.createOrReplaceTempView('users_view');

SELECT *
FROM users_view
LIMIT 100;
```

problem_053

```
-- Problem 053: 053 - Aggregations & GroupBy: Agg challenge
-- Category: Aggregations & GroupBy

-- Register temp view in PySpark:
-- sessions.createOrReplaceTempView('sessions_view');

SELECT *
FROM sessions_view
LIMIT 100;
```

problem_059

```
-- Problem 059: 059 - Joins: Right challenge
-- Category: Joins

-- Register temp view in PySpark:
-- sessions.createOrReplaceTempView('sessions_view');

SELECT *
FROM sessions_view
LIMIT 100;
```

problem_062

```
-- Problem 062: 062 - Joins: Right challenge
-- Category: Joins

-- Register temp view in PySpark:
-- events.createOrReplaceTempView('events_view');

SELECT *
FROM events_view
LIMIT 100;
```

problem_068

```
-- Problem 068: 068 - Window Functions: Dense_rank challenge
-- Category: Window Functions

-- Register temp view in PySpark:
-- users.createOrReplaceTempView('users_view');

SELECT *
FROM users_view
LIMIT 100;
```

problem_069

```
-- Problem 069: 069 - Dates & Timestamps: From_unixtime challenge
-- Category: Dates & Timestamps

-- Register temp view in PySpark:
-- users.createOrReplaceTempView('users_view');

SELECT *
FROM users_view
LIMIT 100;
```

problem_070

```
-- Problem 070: 070 - Spark SQL: Sql queries challenge
-- Category: Spark SQL

-- Register temp view in PySpark:
-- transactions.createOrReplaceTempView('transactions_view');

SELECT *
FROM transactions_view
LIMIT 100;
```

problem_071

```
-- Problem 071: 071 - Strings & Regex: Regexp_replace challenge
-- Category: Strings & Regex

-- Register temp view in PySpark:
-- transactions.createOrReplaceTempView('transactions_view');

SELECT *
FROM transactions_view
LIMIT 100;
```

problem_073

```
-- Problem 073: 073 - Aggregations & GroupBy: Agg challenge
-- Category: Aggregations & GroupBy

-- Register temp view in PySpark:
-- transactions.createOrReplaceTempView('transactions_view');

SELECT *
FROM transactions_view
LIMIT 100;
```

problem_074

```
-- Problem 074: 074 - Spark SQL: Sql queries challenge
-- Category: Spark SQL

-- Register temp view in PySpark:
-- events.createOrReplaceTempView('events_view');

SELECT *
FROM events_view
LIMIT 100;
```

problem_075

```
-- Problem 075: 075 - Joins: Anti challenge
-- Category: Joins

-- Register temp view in PySpark:
-- logs.createOrReplaceTempView('logs_view');

SELECT *
FROM logs_view
LIMIT 100;
```

problem_076

```
-- Problem 076: 076 - Strings & Regex: Concat_ws challenge
-- Category: Strings & Regex

-- Register temp view in PySpark:
-- users.createOrReplaceTempView('users_view');

SELECT *
FROM users_view
LIMIT 100;
```

problem_080

```
-- Problem 080: 080 - Dates & Timestamps: To_date challenge
-- Category: Dates & Timestamps

-- Register temp view in PySpark:
-- events.createOrReplaceTempView('events_view');

SELECT *
FROM events_view
LIMIT 100;
```

problem_083

```
-- Problem 083: 083 - Window Functions: Rank challenge
-- Category: Window Functions

-- Register temp view in PySpark:
-- clicks.createOrReplaceTempView('clicks_view');

SELECT *
FROM clicks_view
LIMIT 100;
```

problem_089

```
-- Problem 089: 089 - Joins: Right challenge
-- Category: Joins

-- Register temp view in PySpark:
-- transactions.createOrReplaceTempView('transactions_view');

SELECT *
FROM transactions_view
LIMIT 100;
```

problem_096

```
-- Problem 096: 096 - Dates & Timestamps: Date_trunc challenge
-- Category: Dates & Timestamps

-- Register temp view in PySpark:
-- clicks.createOrReplaceTempView('clicks_view');

SELECT *
FROM clicks_view
LIMIT 100;
```

problem_098

```
-- Problem 098: 098 - Joins: Left challenge
-- Category: Joins

-- Register temp view in PySpark:
-- clicks.createOrReplaceTempView('clicks_view');

SELECT *
FROM clicks_view
LIMIT 100;
```

problem_101

```
-- Problem 101: 101 - Pivot & Crosstab: Cube challenge
-- Category: Pivot & Crosstab

-- Register temp view in PySpark:
-- events.createOrReplaceTempView('events_view');

SELECT *
FROM events_view
LIMIT 100;
```

problem_105

```
-- Problem 105: 105 - Dates & Timestamps: From_unixtime challenge
-- Category: Dates & Timestamps

-- Register temp view in PySpark:
-- transactions.createOrReplaceTempView('transactions_view');

SELECT *
FROM transactions_view
LIMIT 100;
```

problem_106

```
-- Problem 106: 106 - Dates & Timestamps: From_unixtime challenge
-- Category: Dates & Timestamps

-- Register temp view in PySpark:
-- transactions.createOrReplaceTempView('transactions_view');

SELECT *
FROM transactions_view
LIMIT 100;
```

problem_110

```sql
-- Problem 110: 110 - Joins: Semi challenge
-- Category: Joins

-- Register temp view in PySpark:
-- clicks.createOrReplaceTempView('clicks_view');

SELECT *
FROM clicks_view
LIMIT 100;
```

problem_112

```
-- Problem 112: 112 - Aggregations & GroupBy: Approx_count_distinct challenge
-- Category: Aggregations & GroupBy

-- Register temp view in PySpark:
-- clicks.createOrReplaceTempView('clicks_view');

SELECT *
FROM clicks_view
LIMIT 100;
```

problem_113

```
-- Problem 113: 113 - Pivot & Crosstab: Rollup challenge
-- Category: Pivot & Crosstab

-- Register temp view in PySpark:
-- orders.createOrReplaceTempView('orders_view');

SELECT *
FROM orders_view
LIMIT 100;
```

problem_114

```
-- Problem 114: 114 - Dates & Timestamps: To_date challenge
-- Category: Dates & Timestamps

-- Register temp view in PySpark:
-- events.createOrReplaceTempView('events_view');

SELECT *
FROM events_view
LIMIT 100;
```

problem_116

```
-- Problem 116: 116 - Joins: Semi challenge
-- Category: Joins

-- Register temp view in PySpark:
-- products.createOrReplaceTempView('products_view');

SELECT *
FROM products_view
LIMIT 100;
```

problem_119

```
-- Problem 119: 119 - Joins: Left challenge
-- Category: Joins

-- Register temp view in PySpark:
-- users.createOrReplaceTempView('users_view');

SELECT *
FROM users_view
LIMIT 100;
```

problem_121

```
-- Problem 121: 121 - Pivot & Crosstab: Pivot challenge
-- Category: Pivot & Crosstab

-- Register temp view in PySpark:
-- users.createOrReplaceTempView('users_view');

SELECT *
FROM users_view
LIMIT 100;
```

problem_126

```
-- Problem 126: 126 - Pivot & Crosstab: Rollup challenge
-- Category: Pivot & Crosstab

-- Register temp view in PySpark:
-- events.createOrReplaceTempView('events_view');

SELECT *
FROM events_view
LIMIT 100;
```

problem_127

```
-- Problem 127: 127 - Pivot & Crosstab: Rollup challenge
-- Category: Pivot & Crosstab

-- Register temp view in PySpark:
-- logs.createOrReplaceTempView('logs_view');

SELECT *
FROM logs_view
LIMIT 100;
```

problem_129

```
-- Problem 129: 129 - Dates & Timestamps: Window challenge
-- Category: Dates & Timestamps

-- Register temp view in PySpark:
-- products.createOrReplaceTempView('products_view');

SELECT *
FROM products_view
LIMIT 100;
```

problem_130

```
-- Problem 130: 130 - Pivot & Crosstab: Pivot challenge
-- Category: Pivot & Crosstab

-- Register temp view in PySpark:
-- users.createOrReplaceTempView('users_view');

SELECT *
FROM users_view
LIMIT 100;
```

problem_133

```
-- Problem 133: 133 - Aggregations & GroupBy: Agg challenge
-- Category: Aggregations & GroupBy

-- Register temp view in PySpark:
-- sessions.createOrReplaceTempView('sessions_view');

SELECT *
FROM sessions_view
LIMIT 100;
```

problem_134

```
-- Problem 134: 134 - Pivot & Crosstab: Cube challenge
-- Category: Pivot & Crosstab

-- Register temp view in PySpark:
-- clicks.createOrReplaceTempView('clicks_view');

SELECT *
FROM clicks_view
LIMIT 100;
```

problem_135

```
-- Problem 135: 135 - Spark SQL: Functions in sql challenge
-- Category: Spark SQL

-- Register temp view in PySpark:
-- clicks.createOrReplaceTempView('clicks_view');

SELECT *
FROM clicks_view
LIMIT 100;
```

problem_136

```
-- Problem 136: 136 - Dates & Timestamps: Date_trunc challenge
-- Category: Dates & Timestamps

-- Register temp view in PySpark:
-- products.createOrReplaceTempView('products_view');

SELECT *
FROM products_view
LIMIT 100;
```

problem_137

```
-- Problem 137: 137 - Dates & Timestamps: Date_trunc challenge
-- Category: Dates & Timestamps

-- Register temp view in PySpark:
-- logs.createOrReplaceTempView('logs_view');

SELECT *
FROM logs_view
LIMIT 100;
```

problem_144

```
-- Problem 144: 144 - Strings & Regex: Concat_ws challenge
-- Category: Strings & Regex

-- Register temp view in PySpark:
-- transactions.createOrReplaceTempView('transactions_view');

SELECT *
FROM transactions_view
LIMIT 100;
```

problem_146

```
-- Problem 146: 146 - Window Functions: Row_number challenge
-- Category: Window Functions

-- Register temp view in PySpark:
-- orders.createOrReplaceTempView('orders_view');

SELECT *
FROM orders_view
LIMIT 100;
```

problem_147

```
-- Problem 147: 147 - Window Functions: Rank challenge
-- Category: Window Functions

-- Register temp view in PySpark:
-- orders.createOrReplaceTempView('orders_view');

SELECT *
FROM orders_view
LIMIT 100;
```

problem_149

```
-- Problem 149: 149 - Aggregations & GroupBy: Approx_count_distinct challenge
-- Category: Aggregations & GroupBy

-- Register temp view in PySpark:
-- clicks.createOrReplaceTempView('clicks_view');

SELECT *
FROM clicks_view
LIMIT 100;
```

problem_154

```
-- Problem 154: 154 - Strings & Regex: Split challenge
-- Category: Strings & Regex

-- Register temp view in PySpark:
-- sessions.createOrReplaceTempView('sessions_view');

SELECT *
FROM sessions_view
LIMIT 100;
```

problem_155

```
-- Problem 155: 155 - Pivot & Crosstab: Pivot challenge
-- Category: Pivot & Crosstab

-- Register temp view in PySpark:
-- transactions.createOrReplaceTempView('transactions_view');

SELECT *
FROM transactions_view
LIMIT 100;
```

problem_158

```sql
-- Problem 158: 158 - Joins: Right challenge
-- Category: Joins

-- Register temp view in PySpark:
-- clicks.createOrReplaceTempView('clicks_view');

SELECT *
FROM clicks_view
LIMIT 100;
```

problem_160

```
-- Problem 160: 160 - Spark SQL: Sql queries challenge
-- Category: Spark SQL

-- Register temp view in PySpark:
-- events.createOrReplaceTempView('events_view');

SELECT *
FROM events_view
LIMIT 100;
```

problem_162

```
-- Problem 162: 162 - Dates & Timestamps: Date_trunc challenge
-- Category: Dates & Timestamps

-- Register temp view in PySpark:
-- transactions.createOrReplaceTempView('transactions_view');

SELECT *
FROM transactions_view
LIMIT 100;
```

problem_166

```
-- Problem 166: 166 - Spark SQL: Create temp view challenge
-- Category: Spark SQL

-- Register temp view in PySpark:
-- sessions.createOrReplaceTempView('sessions_view');

SELECT *
FROM sessions_view
LIMIT 100;
```

problem_167

```
-- Problem 167: 167 - Aggregations & GroupBy: Agg challenge
-- Category: Aggregations & GroupBy

-- Register temp view in PySpark:
-- products.createOrReplaceTempView('products_view');

SELECT *
FROM products_view
LIMIT 100;
```

problem_168

```
-- Problem 168: 168 - Spark SQL: Sql queries challenge
-- Category: Spark SQL

-- Register temp view in PySpark:
-- logs.createOrReplaceTempView('logs_view');

SELECT *
FROM logs_view
LIMIT 100;
```

problem_171

```
-- Problem 171: 171 - Dates & Timestamps: From_unixtime challenge
-- Category: Dates & Timestamps

-- Register temp view in PySpark:
-- clicks.createOrReplaceTempView('clicks_view');

SELECT *
FROM clicks_view
LIMIT 100;
```

problem_175

```
-- Problem 175: 175 - Strings & Regex: Concat_ws challenge
-- Category: Strings & Regex

-- Register temp view in PySpark:
-- users.createOrReplaceTempView('users_view');

SELECT *
FROM users_view
LIMIT 100;
```

problem_178

```
-- Problem 178: 178 - Window Functions: Lag challenge
-- Category: Window Functions

-- Register temp view in PySpark:
-- products.createOrReplaceTempView('products_view');

SELECT *
FROM products_view
LIMIT 100;
```

problem_179

```
-- Problem 179: 179 - Aggregations & GroupBy: Groupby challenge
-- Category: Aggregations & GroupBy

-- Register temp view in PySpark:
-- logs.createOrReplaceTempView('logs_view');

SELECT *
FROM logs_view
LIMIT 100;
```

problem_184

```
-- Problem 184: 184 - Joins: Inner challenge
-- Category: Joins

-- Register temp view in PySpark:
-- orders.createOrReplaceTempView('orders_view');

SELECT *
FROM orders_view
LIMIT 100;
```

problem_185

```
-- Problem 185: 185 - Dates & Timestamps: From_unixtime challenge
-- Category: Dates & Timestamps

-- Register temp view in PySpark:
-- sessions.createOrReplaceTempView('sessions_view');

SELECT *
FROM sessions_view
LIMIT 100;
```

problem_187

```
-- Problem 187: 187 - Spark SQL: Sql queries challenge
-- Category: Spark SQL

-- Register temp view in PySpark:
-- sessions.createOrReplaceTempView('sessions_view');

SELECT *
FROM sessions_view
LIMIT 100;
```

problem_188

```
-- Problem 188: 188 - Spark SQL: Functions in sql challenge
-- Category: Spark SQL

-- Register temp view in PySpark:
-- transactions.createOrReplaceTempView('transactions_view');

SELECT *
FROM transactions_view
LIMIT 100;
```

problem_189

```
-- Problem 189: 189 - Joins: Inner challenge
-- Category: Joins

-- Register temp view in PySpark:
-- events.createOrReplaceTempView('events_view');

SELECT *
FROM events_view
LIMIT 100;
```

problem_193

```
-- Problem 193: 193 - Spark SQL: Create temp view challenge
-- Category: Spark SQL

-- Register temp view in PySpark:
-- products.createOrReplaceTempView('products_view');

SELECT *
FROM products_view
LIMIT 100;
```

problem_199

```
-- Problem 199: 199 - Strings & Regex: Regexp_replace challenge
-- Category: Strings & Regex

-- Register temp view in PySpark:
-- orders.createOrReplaceTempView('orders_view');

SELECT *
FROM orders_view
LIMIT 100;
```

problem_200

```
-- Problem 200: 200 - Pivot & Crosstab: Cube challenge
-- Category: Pivot & Crosstab

-- Register temp view in PySpark:
-- clicks.createOrReplaceTempView('clicks_view');

SELECT *
FROM clicks_view
LIMIT 100;
```