

# EXAMPLE - SUMIF and COUNTDISTINCTIF Functions

This example illustrates how you can use conditional calculation functions.

## Functions:

Item	Description
SUMIF Function	Generates the sum of rows in each group that meet a specific condition.
COUNTDISTINCTIF Function	Generates the count of distinct non-null values for rows in each group that meet a specific condition.

## Source:

Here is some example polling data across 16 precincts in 8 cities across 4 counties, where registrations have been invalidated at the polling station, preventing voters from voting. Precincts where this issue has occurred previously have been added to a watch list (`precinctWatchList`).

totalReg	invalidReg	precinctWatchList	precinctId	cityId	countyId
731	24	y	1	1	1
743	29	y	2	1	1
874	0		3	2	1
983	0		4	2	1
622	29		5	3	2
693	0		6	3	2
775	37	y	7	4	2
1025	49	y	8	4	2
787	13		9	5	3
342	0		10	5	3
342	39	y	11	6	3
387	28	y	12	6	3
582	59		13	7	4
244	0		14	7	4
940	6	y	15	8	4
901	4	y	16	8	4

## Transformation:

First, you want to sum up the invalid registrations (`invalidReg`) for precincts that are already on the watchlist (`precinctWatchList = y`). These sums are grouped by city, which can span multiple precincts:

<b>Transformation Name</b>	New formula
<b>Parameter: Formula type</b>	Single row formula
<b>Parameter: Formula</b>	<code>SUMIF(invalidReg, precinctWatchList == "y")</code>

<b>Parameter: Group rows by</b>	cityId
<b>Parameter: New column name</b>	'invalidRegbyCityId'

The `invalidRegbyCityId` column contains invalid registrations across the entire city.

Now, at the county level, you want to identify the number of precincts that were on the watch list and were part of a city-wide registration problem.

In the following step, the number of cities in each count are counted where invalid registrations within a city is greater than 60.

- This step creates a pivot aggregation.

<b>Transformation Name</b>	Pivot columns
<b>Parameter: Row labels</b>	countyId
<b>Parameter: Values</b>	COUNTDISTINCTIF(precinctId, invalidRegbyCityId > 60)
<b>Parameter: Max number of columns to create</b>	1

### Results:

countyId	countdistinctif_precinctId
1	0
2	2
3	2
4	0

The voting officials in counties 2 and 3 should investigate their precinct registration issues.