

Dashboard

Guided Project: Clean And Analyze Employee Exit Surveys

LEARNMISSION REFERENCE

Learn

In the last screen, we created a new `institute_service` column that we'll use to analyze survey respondents according to their length of employment. Next, we'll identify any employees who resigned because they were dissatisfied.

Below are the columns we'll use to categorize employees as "dissatisfied" from each dataframe. If you disagree, feel free to modify them! Just make sure you explain *why* you made that decision.

- tafe_survey_updated:
 - Contributing Factors. Dissatisfaction
 - Contributing Factors. Job Dissatisfaction
- detesurveyupdated:
 - job_dissatisfaction
 - dissatisfaction_with_the_department
 - physical_work_environment
 - lack_of_recognition
 - lack_of_job_security
 - work_location
 - employment_conditions
 - work_life_balance
 - workload

If the employee indicated any of the factors above caused them to resign, we'll mark them as `dissatisfied` in a new column.

To create the new column, we'll do the following:

- Convert the values in the `'Contributing Factors. Dissatisfaction'` and `'Contributing Factors. Job Dissatisfaction'` columns in the `tafe_resignations` dataframe to `True`, `False`, or `NaN` values

Instructions

`dissatisfied`. To accomplish this, we'll use the `DataFrame.any()` method to do the following:

- Return `True` if any element in the selected columns above is `True`
- Return `False` if none of the elements in the selected columns above is `True`
- Return `NaN` if the value is `NaN`

True	False	True
False	False	True
False	False	False
NaN	NaN	NaN

True
True
False
NaN

Here's the syntax we can use:

```
df.any(axis=1, skipna=False)
```

After our changes, the new `dissatisfied` column will contain just the following values:

- `True`: Indicates a person resigned because they were dissatisfied with the job
- `False`: indicates a person resigned because of a reason other than dissatisfaction with the job

Instructions

Instructions

- Use the `Series.value_counts()` method to view the values in the `'Contributing Factors. Dissatisfaction'` and `'Contributing Factors. Job Dissatisfaction'` in the `tafe_resignations` dataframe.
- Update the values in the `'Contributing Factors. Dissatisfaction'` and `'Contributing Factors. Job Dissatisfaction'` in the `tafe_resignations` dataframe so that each contains only `True`, `False`, or `NaN` values.
 - Write a function named `update_vals` that makes the following changes:
 - If the value is `NaN`, return `np.nan`. You can use the following criteria to check that a value is `NaN`: `pd.isnull(val)`.
 - If the value is `'-'`, return `False`.
 - For any other value, return `True`.
 - Use the `DataFrame.applymap()` method to apply the function above to the `'Contributing Factors. Dissatisfaction'` and `'Contributing Factors. Job Dissatisfaction'` in the `tafe_resignations` dataframe.
 - Remember that we need to pass the `update_vals` function into the `df.applymap()` method *without* parentheses.
 - Use the `df.any()` method as described above to create a `dissatisfied` column in BOTH the `tafe_resignations` and `dete_resignations` dataframes.
 - Use the `df.copy()` method to create a copy of the results and avoid the SettingWithCopy Warning. Assign the results to `dete_resignations_up` and `tafe_resignations_up`.
 - Write a markdown paragraph explaining the changes you made and why.

Get Help

Basics.ipynbRefreshDownload

jupyter Basics (autosaved)

FileEditViewInsertCellKernelHelpPython 3

maximum of 49 years and a mean of 10 years.

Identifying Dissatisfied Employees

We'll identify employees who left because of dissatisfaction. Below are the columns I'll use to categorize employees as 'dissatisfied' from each dataframe.

- dete_resignations:
 - job_dissatisfaction
 - dissatisfaction_with_the_department
 - physical_work_environment
 - lack_of_recognition
 - lack_of_job_security
 - work_location
 - employment_conditions
 - work_life_balance
 - workload
- tafe_resignations:
 - Contributing Factors. Dissatisfaction
 - Contributing Factors. Job Dissatisfaction
 - Contributing Factors. Interpersonal Conflict

If an employee indicated that any of the factors above caused him/her to resign, I'll mark them as dissatisfied in a new column.

In []: