

Nurse Staffing Recommendations

Matthew Bain

2024-03-22

abstract I investigate nurse staffing data to provide informed recommendations to a medical staffing organization.

Table of contents

Imports	1
The dataset	1
Some GT examples	12
SQL	14
Bibliography	14

[...]

Imports

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

from great_tables import GT
from pandas.plotting import scatter_matrix

from src.stylesheet import customize_plots
from src.inspection import make_df, display
```

The dataset

We begin by exploring the data to get to know the features and patterns on which we will base our analysis.

```
if 'data' not in locals():
    data = pd.read_csv(
        "../data/raw/PBJ_Daily_Nurse_Staffing_Q1_2024.zip",
        encoding='ISO-8859-1'
```

```
)  
else:  
    print("data loaded.")
```

```
data.sample(5)
```

PROVNUM	PROVCITY	STATE	COUNTY	QW	Work-MF	RN	Hrs	PC	Nrs	Crs	CH	Hrs	Wk	Ad	Ampe	ctr
NAME	TY	TY	-	Date	Scen	DON	N_ctr	NA_e	NA_ctr	NA	NA	NA	NA	NA	NA	NA
	NAME	FIPS				sus						ttm	ner	tp	ctr	
115513	TWIN	LA	MAGA	Ba-	52024	102401072	8.0	...	36.2	146.4	146.4	0.0	0.0	0.0	0.00	0.00
	OAKS		con													
233053	CON-															
	VA-															
	LES-															
	CENT															
	CEN-															
	TER															
46A01A	MAPLE	LA	TUT	Salt	352024	102402034	0.0	...	0.00	47.16	47.16	0.0	0.0	0.0	0.00	0.00
	RIDGE		Lake													
	RE-CITY															
	HA-															
1109690	BIL-															
	ITA-															
	TION															
	AND															
	NURS-															
	ING															
44542A	HOMER	TN	Shelby	152024	102402258	8.0	...	47.98	10.59	99.00	11.5	0.0	0.0	0.0	0.00	0.00
1071036	BRIGHT															
	GLADE															
36623A	OLIVE	OH	Mer-	102024	102403177	0.0	...	0.00	38.25	38.25	0.0	0.0	0.0	0.0	0.00	0.00
	NIAEORD		cer													
	NURS-															
	ING															
902705	CEN-															
	TER															
	OF															
	ROCK-															
	FORD															
37512A	SEFAHL	OK	Chero-	212024	102401319	8.0	...	0.00	80.92	80.92	0.0	0.0	0.0	0.0	14.83	14.83
	QUOYAH		kee													
	POINTE															
924408	SKILLED															
	NURS-															
	ING															
	AND															
	THER-															
	APY															

```
# TODO: pivot on day

data_pivoted = data.pivot_table(
    index="STATE",
    columns="WorkDate",
    values="Hrs_RN",
    aggfunc='mean'
)

# Resetting the index for easier column access
# data_pivoted.reset_index(inplace=True)
data_pivoted
```



```
data_pivoted.iloc[:, 1:]
```



```

for ax in axs.flatten():
    ax.tick_params(axis='both', which='both', length=3.5)

# save_fig("scatter_matrix_plot")

plt.show()

```

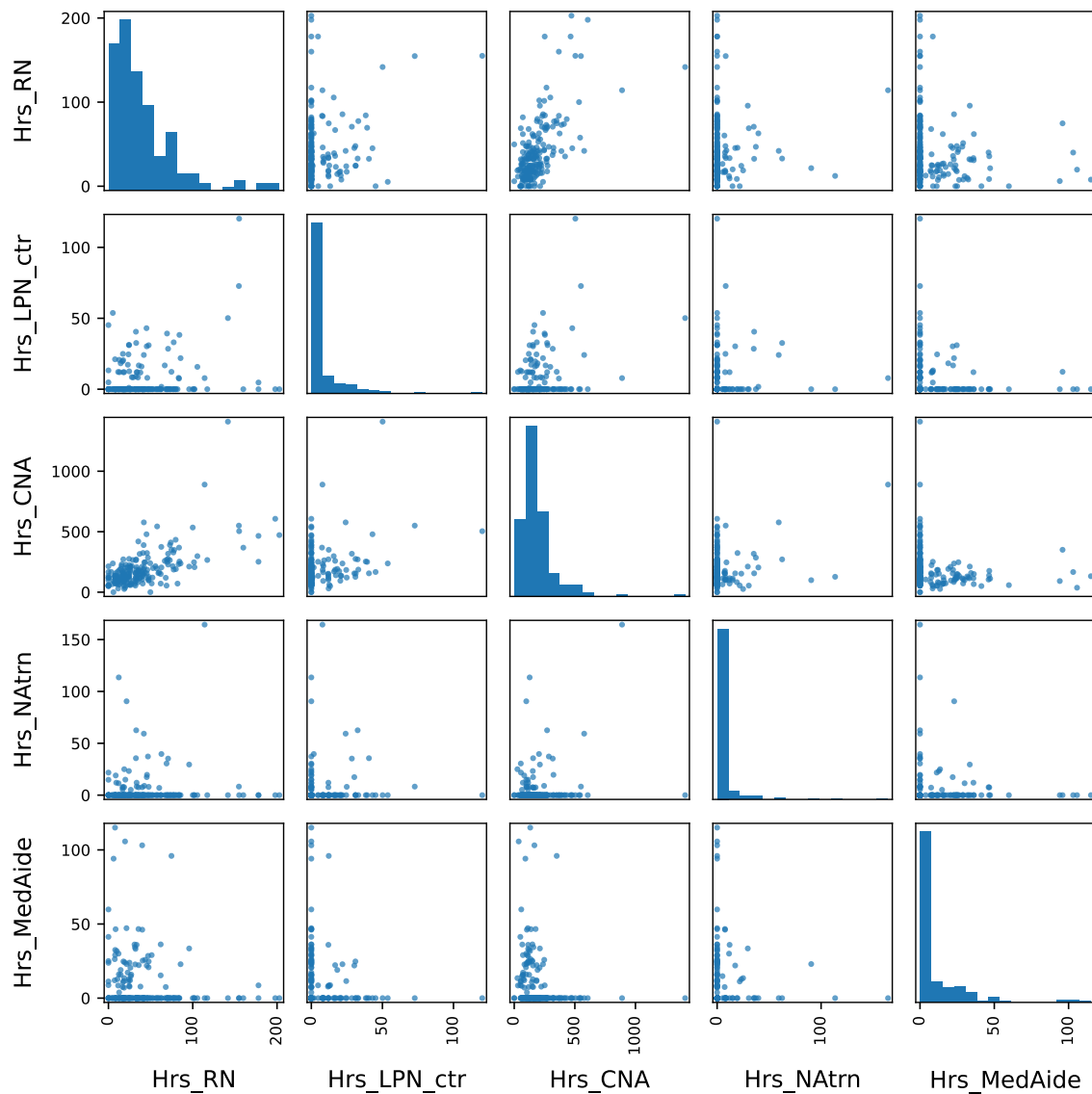


Figure 1: Scatter matrix of nursing worker working hours

```

from great_tables import GT

df = data.loc[150000:, [

```

```
"STATE",  
"COUNTY_NAME", "COUNTY_FIPS",  
"CITY",  
"PROVNAME", "PROVNUM",  
# "MDScensus"  
]].value_counts().reset_index()  
GT(df.head(n=10))
```

STATE	COUNTY_- NAME	COUNTY_- FIPS	CITY	PROV- NAME	PROVNUM	count
CA	Alameda	1	ALAMEDA	ALAMEDA HEALTH- CARE & WELL- NESS CEN- TER	555486	91
OH	Mahoning	99	AUSTIN- TOWN	AVEN- TURA AT HUMILITY HOUSE	366186	91
OH	Lucas	95	TOLEDO	OHIO LIV- ING SWAN CREEK	365996	91
OH	Lucas	95	TOLEDO	OTTER- BEIN SUNSET HOUSE	366148	91
OH	Lucas	95	TOLEDO	PARK TER- RACE NURSING AND RE- HABILITA- TION CEN- TER	365339	91
OH	Lucas	95	TOLEDO	POINT PLACE HEALTH- CARE AND REHABIL- ITATION CENTER	366039	91
OH	Lucas	95	WATER- VILLE	ASTORIA PLACE OF AYDEN WATER- VILLE	365747	91
OH	Madison	95	WATER- LONDON MOUSE	MIDWEST HEALTH- CARE OF COUNTESS WATER- VILLE	365747	91

Some GT examples

```
from typing import Any
from IPython.display import display as ipy_display, HTML
import numpy as np

def display2(
    *args,
    globs: dict[str, Any] | None = None,
    bold: bool = True,
    width: str = "400px" # Fixed width for each block
) -> None:
    """
    Display an informative representation of multiple objects side-by-side in
    Jupyter.

    Parameters
    -----
    *args : tuple
        Tuple of expressions to evaluate and display.
    globs : dict[str, Any], default=None
        Global namespace, to give eval() access to nonlocals passed by name.
    bold : bool, default=True
        Option to enable/disable string styling.
    width : str, default="400px"
        Fixed width for each displayed block in the Jupyter notebook.

    Warnings
    -----
    This function uses `eval()` to render expressions it receives
    as strings. Access to variables in the global namespace is controlled
    by `globs`. Take care to only pass trusted expressions to the function.
    """

    if globs is None:
        globs = {}

    outputs = []
    for arg in args:
        name = f"<b>{arg}</b>" if bold else arg
        value = np.round(eval(arg, globs), 2)
        shape = np.shape(value)
        content = f"<div style='width:{width}; padding:10px; float:left;'><pre>{name}\n--- {repr(shape)} ---\n{repr(value)}</pre></div>"
        outputs.append(content)

    # Clearfix for layout
    clearfix = "<div style='clear: both;'></div>"
```

```
# Display the HTML content in Jupyter
html_output = ''.join(outputs) + clearfix
ipy_display(HTML(html_output))
```

```
return None
```

```
A = np.array([[1, 3], [2, 4]])
x = np.array([[0, 1]])

display2(
    "A", "x.T", "np.dot(A, x.T)", globs=globals(), bold=True, width="100px"
)
```

```
<IPython.core.display.HTML object>
```

```
display2(
    "data['STATE'].value_counts()",
    "data['COUNTY_NAME'].value_counts()",
    "data['CITY'].value_counts()",
    "data['PROVNAME'].value_counts()",
    "data['MDScensus'].value_counts()",
    width="340px",
    globs=globals()
)
```

```
<IPython.core.display.HTML object>
```

```
data[["CY_Qtr", "WorkDate", "MDScensus"]]
```

	CY_Qtr	WorkDate	MDScensus
0	2024Q1	20240101	50
1	2024Q1	20240102	49
2	2024Q1	20240103	49
3	2024Q1	20240104	50
4	2024Q1	20240105	51
...
1330961	2024Q1	20240327	81
1330962	2024Q1	20240328	83
1330963	2024Q1	20240329	85
1330964	2024Q1	20240330	82
1330965	2024Q1	20240331	82

SQL

Bibliography