

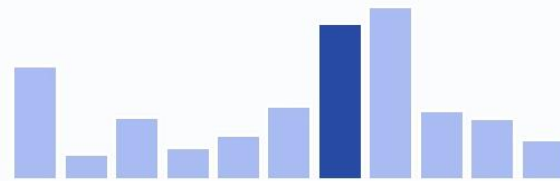
Módulo 2 – Obtención y Preparación de Datos

# Actividad: Limpieza Sueldos San Francisco

Ciencia de Datos

# Actividad: “Caso Limpieza Sueldos San Francisco”

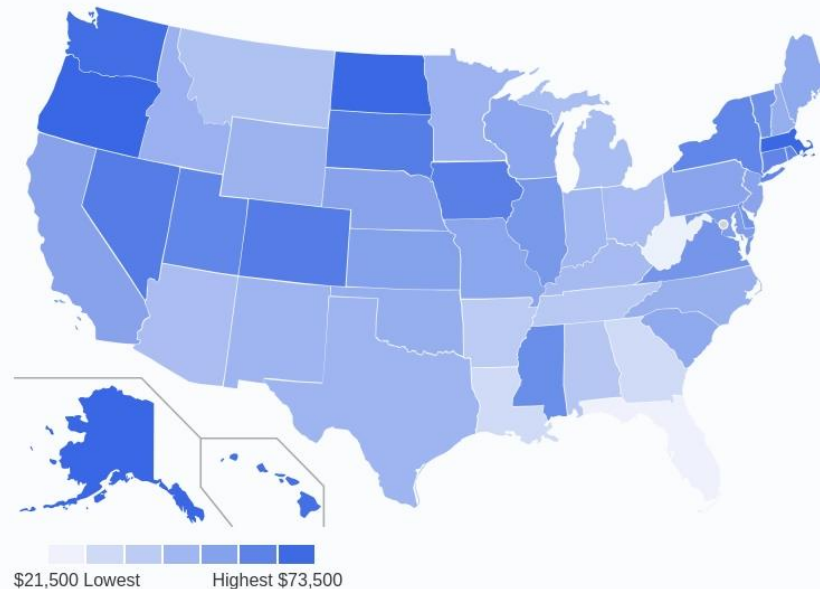
## San Francisco Bay Area Salary in the United States



National Average

\$50,117 /year  
\$24 /hour

	Annual Salary	Monthly Pay	Weekly Pay	Hourly Wage
Top Earners	\$68,500	\$5,708	\$1,317	\$33
75th Percentile	\$59,000	\$4,916	\$1,134	\$28
Average	\$50,117	\$4,176	\$963	\$24
25th Percentile	\$38,500	\$3,208	\$740	\$19



# Caso Limpieza Sueldos San Francisco

Después del análisis de datos que usted realizó sobre la nómina de sueldos de San Francisco, el alcalde notó los errores de datos que existían y los problemas que acarrearán a la gestión. Por este motivo, decidió contratarlo para que haga un análisis de la calidad de datos y realice la limpieza de datos a efecto de contar con información confiable para la toma de decisiones.

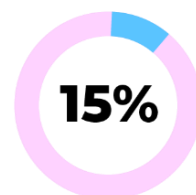
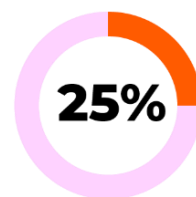
	Id	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBenefits	Year	Notes	Agency	Status
0	1	NATHANIEL FORD	GENERAL MANAGER-METROPOLITAN TRANSIT AUTHORITY	167411.18	0.00	400184.25	NaN	567595.43	567595.43	2011	NaN	San Francisco	NaN
1	2	GARY JIMENEZ	CAPTAIN III (POLICE DEPARTMENT)	155966.02	245131.88	137811.38	NaN	538909.28	538909.28	2011	NaN	San Francisco	NaN
2	3	ALBERT PARDINI	CAPTAIN III (POLICE DEPARTMENT)	212739.13	106088.18	16452.60	NaN	335279.91	335279.91	2011	NaN	San Francisco	NaN
3	4	CHRISTOPHER CHONG	WIRE ROPE CABLE MAINTENANCE MECHANIC	77916.00	56120.71	198306.90	NaN	332343.61	332343.61	2011	NaN	San Francisco	NaN



# Caso Limpieza Sueldos San Francisco

Como entregable, se solicita un notebook ordenado, documentado y reproducible en donde:

- Realice un análisis de valores nulos en las columnas **BasePay**, **OverTimePay**, **OtherPay**, **Benefits**, **TotalPay** y **TotalPayBenefits**
- Realice un análisis de outliers en la columna BasePay
- Realice la limpieza de los datos de acuerdo con los criterios aprendidos (dejar, imputar, eliminar).
- Genere un nuevo archivo Excel con los datos limpios.



¡Éxito!