

Exploring Healthcare Data, Part 2

Testo originario: Meredith Griffith

Traduzione a cura di GMSL

Nella [prima parte](#) di questo articolo, abbiamo visto un case study nel quale lo staff di un ospedale ha usato il test del tampone ATP per verificare la presenza del batterio su 8 superfici presenti in 10 stanze dell'ospedale all'interno di 5 reparti. Le misurazioni di ATP sotto le 400 unità superano il test del tampone, mentre misure superiori o uguali a 400 unità lo falliscono e richiedono ulteriori approfondimenti.



Abbiamo proposto due consigli sull'esplorare e sul visualizzare i dati usando grafici, il brushing e la formattazione condizionata.

1. Valutare la forma dei vostri dati
2. Identificare e analizzare gli outliers

Effettuando queste esplorazioni preliminari sui dati del test tampone, abbiamo scoperto che la misurazione della media ATP non sarebbe stata sufficiente da testare nel caso in cui le superfici avessero mostrato differenze statisticamente significative nei livelli di contaminazione. Questo era dovuto al fatto che i dati risultavano inclinati dagli outliers estremi.

Abbiamo poi identificato dove questi inusuali alti livelli di ATP sono stati scoperti nell'ospedale.

Queste scoperte forniscono delle informazioni utilizzabili in maniera appropriata per focalizzare gli sforzi e per migliorare i processi in particolari stanze di ospedale, reparti o nelle superfici all'interno di queste stanze.

Ora che abbiamo visto quanto una semplice esplorazione e degli strumenti di visualizzazione possono rivelare, andiamo a verificare altri tre strumenti che vi aiuteranno a esplorare i vostri dati sanitari per progettare idee attuabili.

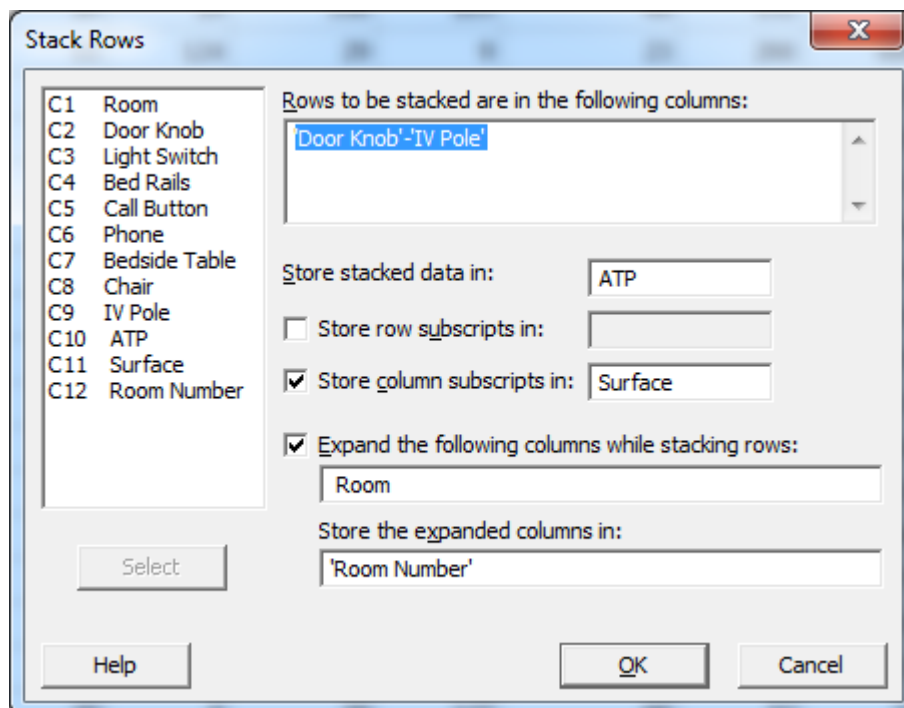
Se volete seguirci passo passo e non avete scaricato i dati dal primo articolo, [potete scaricare e esplorare i dati da soli](#). Se non avete ancora Minitab 17 potete scaricare la versione trial [qui](#).

Consiglio n°3: Manipolare i dati

I dati riguardanti il test tampone che lo staff dell'ospedale ha raccolto, sono suddivisi in più colonne, questo significa semplicemente che tutte le risposte delle misurazioni sono contenute in colonne multiple, anziché raccolte in un'unica colonna. Per aggiungere visibilità dei dati e un'analisi più formale avete bisogno di riconfigurare o manipolare i dati. Possiamo farlo mettendo in un'unica colonna tutte le righe (stacking rows).

Il worksheet ATP Stacked.MTW, all'interno del progetto Minitab scaricabile più in alto, ha già i dati riformati per voi. Ma potete manipolarli voi stessi usando il worksheet ATP Unstacked.MTW

Cliccate su **Data > Stack > Rows** e completate il dialogo come mostrato sotto



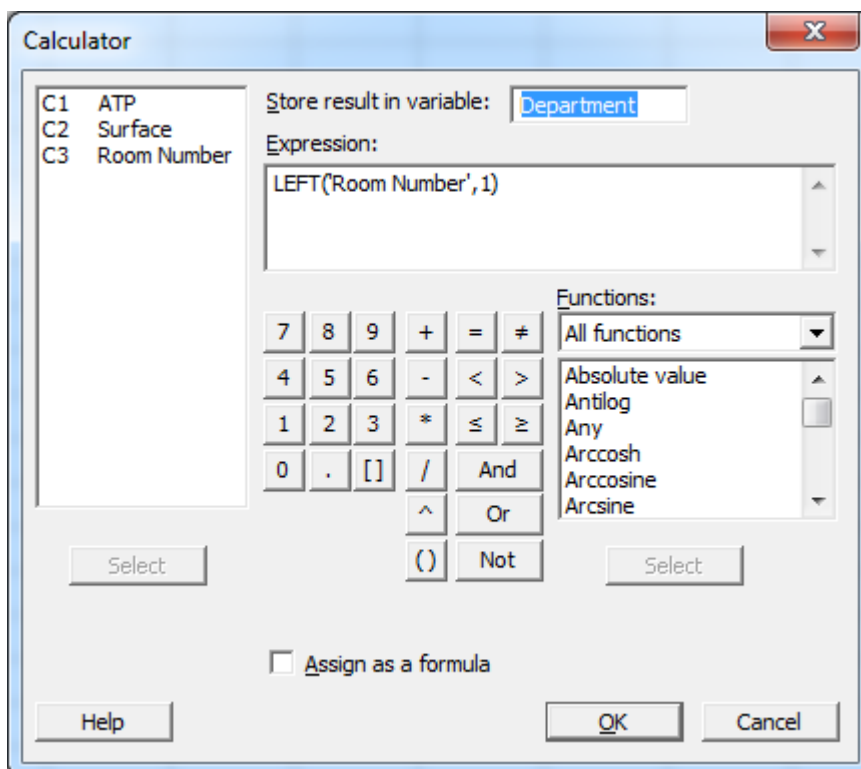
Allocare tutte le righe dei dati e memorizzare la colonna nominata in una colonna separata che farà risultare tutte le misurazioni dell'ATP in un'unica colonna, separata, contenente le categorie per superfici e un'altra colonna contenente il valore "Room Number"

Con i dati *stacked* avete correttamente impostato Minitab per effettuare un'analisi formale : questo è un passo importante tanto quanto lavorare con i vostri dati, in quanto la maggior parte delle analisi di Minitab richiede colonne di dati accatastati. Noi non affronteremo un'analisi formale, ma vi assicuriamo che i dati sono impostati per poterla effettuare.

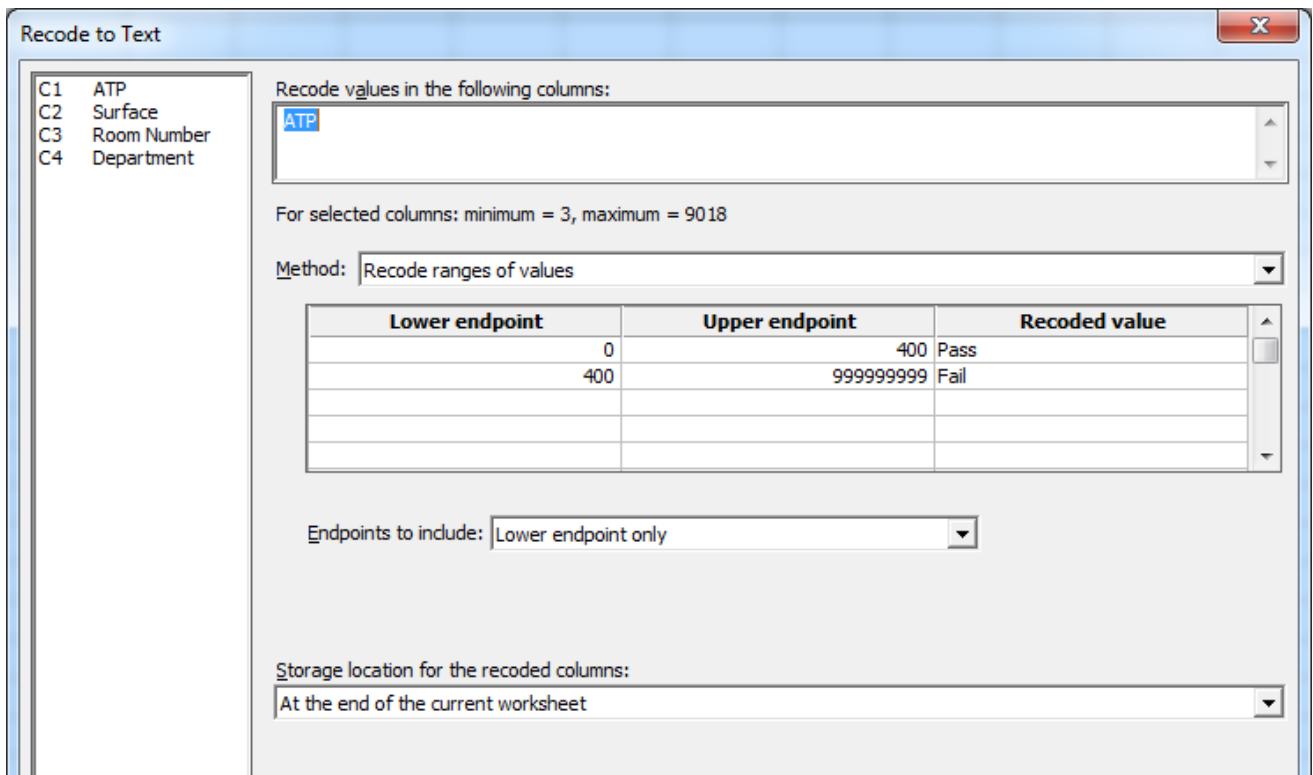
Consiglio N°4: Estrarre le informazioni dal vostro data set originale

Una volta che i vostri dati sono stati accatastati, potete usare altre funzioni disponibili in **Calc > Calculator** e **Data > Recode** per sfruttare le informazioni intrinseche nei vostri dati originali per creare nuove variabili da esplorare e analizzare.

Per esempio sappiamo che il primo carattere di ogni stanza indica il reparto. Potete usare la funzione “Left” in **Calc > Calculator** per estrarre il carattere più a sinistra dalla colonna “Room” e memorizzare il risultato in una nuova colonna denominata “Department” (Reparto). Potete anche farlo compilando nel **Calculator** il dialogo come sotto riportato

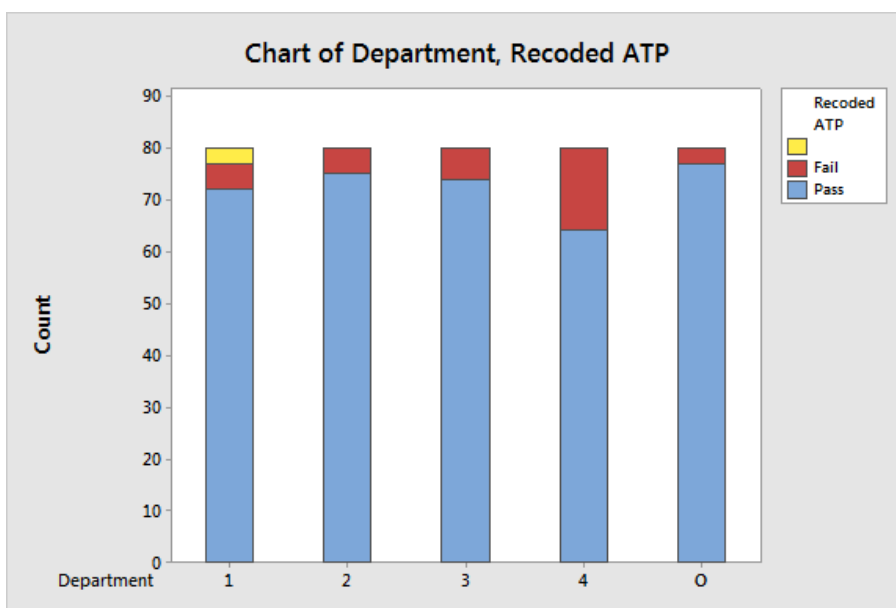


Sapete inoltre che la misurazione sotto il valore di 400 passa il test tampone. Associare i range dei valori ATP a un testo per indicare quale valore passa il test (“Pass”) e quale lo fallisce (“Fail”) può essere utile nel momento in cui andiamo a visualizzare i dati.



Infine, potete usare questi nuovi dati estratti per creare una stacked bar chart che mostra i conti delle misurazioni che hanno fallito, quelle che hanno passato il test o quelle mancante dal test tampone dell'ATP lungo i reparti . Usando il foglio di lavoro **ATP Stacked.MTW** cliccate su **Graph > Bar Chart > Stack**. **Verificate che le barre rappresentate mostrino i valori all'interno della selezione Counts of unique values**. **Cliccate OK**. **Selezionate Select Department e Recoded ATP** come variabili categoriche e cliccate **OK**

Minitab produrrà il seguente grafico:



Il bar chart mostra che:

- il reparto numero 4 ha il valore più alto di ATP che ha fallito il test del tampone
- Lo staff delle pulizie dovrebbe focalizzare inizialmente gli sforzi nel reparto numero 4 in modo che l'analisi delle procedure di pulizia delle stanze vada avanti.

Consiglio n°5: Ottenere statistiche che descrivono i dati

Ora che abbiamo manipolato i dati in un modo che ci prepara per analisi più formale, identifichiamo quale reparto contiene le superfici più contaminate e compariamola con i valori di ogni reparto che hanno passato o fallito il test del tampone ATP, così facendo possiamo disporre delle analisi di statistica descrittiva per dare un'idea di quanto la media o la mediana dei livelli dei batteri differiscono o variano lungo le superfici o a seconda dei reparti.

Usando il worksheet ATP Stacked.MTW, cliccate su **Stat > Basic Statistics > Display Descriptive Statistics**. Inserite ATP come Variabile, Reparto (“Department”) come il By variable e cliccate su OK. Premete CTRL+E per reinserire la finestra di dialogo per la statistica descrittiva e modificare i reparti “Department” con le superfici “Surface” come il By variable. Cliccate OK. Verrà fuori il seguente Output nella finestra di sessione

Descriptive Statistics: ATP

Variable	Department	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
ATP	1	77	3	104.0	15.3	134.6	5.0	24.0	56.0	131.0	689.0
	2	80	0	247	101	901	3	30	59	129	5780
	3	80	0	245	112	1004	7	36	90	170	9018
	4	80	0	245.5	32.8	293.4	15.0	64.5	137.5	347.5	2010.0
	0	80	0	95.0	12.5	111.9	4.0	27.8	52.5	106.0	536.0

Descriptive Statistics: ATP

Variable	Surface	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
ATP	Bed Rails	50	0	237	115	813	8	38	90	163	5780
	Bedside Table	50	0	268	116	818	6	31	72	179	5754
	Call Button	50	0	331	178	1262	3	40	86	283	9018
	Chair	50	0	128.0	17.9	126.9	11.0	50.8	79.5	155.3	631.0
	Door Knob	47	3	97.4	17.4	119.0	3.0	22.0	57.0	122.0	615.0
	IV Pole	50	0	127.4	21.3	150.4	10.0	31.8	65.0	144.8	689.0
	Light Switch	50	0	101.1	18.2	128.9	5.0	23.3	51.0	111.3	536.0
	Phone	50	0	208.0	47.8	338.3	6.0	35.0	89.5	205.8	2010.0

La statistica descrittiva mostra informazioni utili:

- Queste statistiche consentono un facile confronto di media e misurazioni ATP mediana nonché la variazione delle misurazioni ATP, sia per dipartimento che per superficie
- Notare che i valori della media ATP sono più alte dei valori della mediana ATP per entrambi i blocchi di statistica descrittiva. Questo è perché i dati sono spostati sulla destra. Alcune analisi che partono dal presupposto che i dati siano distribuiti normalmente – come ad esempio il t-test compara le

medie – non sono il miglior strumento per analizzare questi tipi di dati. Comparare le mediane potrebbe offrire un quadro più chiaro.

- Entrambi i blocchi di statistica descrittiva evidenziano su quali reparti e superfici focalizzarsi per analizzare e migliorare gli sforzi riguardanti il processo. Ad esempio il reparto 4 ha la più alta mediana di presenza di ATP, mentre le variabili Bed Rails, Phone, e Call Button – i punti di contatto più vicini al paziente malato in un letto di ospedale – sembrano essere le più problematiche superfici da sanificate. Gli sforzi di miglioramento di processo possono iniziare con queste informazioni.

Cos'altro potete fare con i vostri dati?

Quello che avete visto in questa seconda parte di articolo è solo l'inizio. Considerate quanto di questa esplorazione iniziale sia implementabile! Avendo questa base per visualizzare e manipolare i vostri dati, sarete pronti ad investigare e testare le cause per predisporre delle analisi più performanti che producano risultati affidabili.