7) Resuma sus resultados. En particular, explique las implicaciones de la relación entre los ingresos y el número de visitas al sitio web. Siéntase libre de examinar cualquier otro par de variables que considere importante.

8) El departamento de control de calidad está interesado en modelar datos críticos para su negocio. Por ejemplo, si los datos de una variable en particular parecen aproximarse razonablemente mediante una distribución normal, con una media y una desviación estándar predecibles, los valores futuros de esa variable pueden estimarse razonablemente. El propósito del siguiente ejercicio es continuar con este proceso de modelado.

La hoja de cálculo de libras vendidas contiene las libras de material vendidas por semana desde el 3 de enero de 2005 hasta la semana del 19 de julio de 2010.

a) Determine los siguientes valores de resumen para estos datos: media, mediana, desviación estándar, mínimo y máximo.

b) Cree un histograma con los datos de libras de material vendido

Consejo de Excel: Histogramas. Puede encontrar un buen tutorial sobre cómo crear histogramas en Excel en http://www.stat.ncsu.edu/people/reiland/courses/st101/1-4\_excelhisto/sld001.html.

*Algunas sugerencias: Utilice el enfoque dado en la diapositiva cinco como una estimación del ancho de la clase; puede que desee experimentar con valores más pequeños o más grandes. (Otra regla general para determinar el número de intervalos es comenzar con la raíz cuadrada del número de observaciones y ajustar a partir de ahí).*

*Este tutorial utiliza Excel 2003. Utilice las siguientes modificaciones para Excel 2007/2010:*

*En la diapositiva siete, elija Datos/Análisis de datos (no Herramientas/Análisis de datos).*

*Reemplace los pasos de las diapositivas 13 a 15 con los siguientes (para reducir los espacios entre las barras a cero): haga clic con el botón derecho en una barra, elija Formato de serie de datos, establezca Ancho de espacio en cero y haga clic en Cerrar*

*Las sugerencias de "ajuste fino" de las diapositivas 18 en adelante son opcionales. Sin embargo, en Excel 2007/2010 debe delinear cada barra del histograma en negro: haga clic con el botón derecho en una barra, elija Formato de serie de datos, Color del borde, elija Línea sólida, elija negro como color y luego haga clic en Cerrar.*

*Otro buen tutorial (en video) que usa Excel 2010 se encuentra en: http://www.youtube.com/watch?v=RyxPp22x9PU..*

*Tenga en cuenta que este enfoque utiliza dos pasos para definir mejor los anchos de intervalo que aparecen en el histograma.*

c) Describa el histograma. ¿Tiene forma de campana?

Generalmente, simplemente examinar el histograma tiene limitaciones para juzgar qué tan bien un conjunto de datos sigue la distribución normal. (Tenga en cuenta, por ejemplo, que si bien todas las distribuciones normales son unimodales y simétricas, no todas las distribuciones unimodales y simétricas son normales). Un enfoque adicional es ver qué tan bien los datos siguen la regla empírica

d) Determine qué tan bien estos datos siguen la regla empírica completando la siguiente tabla.

A table with text on it

Description automatically generated

El número teórico de observaciones se refiere al número de observaciones en cada intervalo si los datos siguieran exactamente la regla empírica. Por lo tanto, calcule el número teórico de observaciones para cada intervalo multiplicando el número de observaciones en el conjunto de datos de libras vendidas por el porcentaje teórico de datos en ese intervalo.

Para determinar el número real de observaciones, deberá contar el número de observaciones en cada intervalo. La mejor manera de hacerlo es crear una tercera columna en el

Hoja de cálculo de libras vendidas que determina la puntuación z para cada observación. Recuerde que la puntuación z se calcula como:

A black and blue text

Description automatically generated

Por lo tanto, la puntuación z indica el número de desviaciones estándar a las que una observación se aleja de la media. (Por ejemplo, si la media y la desviación estándar de un conjunto de datos son 25 y 5 respectivamente, entonces una observación de 35 está a (35 - 25)/5 = 2 desviaciones estándar de la media). Una vez que haya calculado las puntuaciones z para los datos de libras vendidas, ordene los datos por las puntuaciones z. Entonces será sencillo determinar el número real de observaciones dentro de 1, 2 o 3 desviaciones estándar de la media.

*Consejo de Excel: Conteo. Si ha ordenado sus datos por puntuación z, puede contar cuáles*

*Las observaciones están en cada categoría directamente. Alternativamente, puede usar la función de Excel =countif() para contar todas las entradas mayores que (>), menores que (<), mayores o iguales que (>=) o menores o iguales que (<) algún valor. Por ejemplo, si desea determinar el número de observaciones en un rango que son menores o iguales a 1, puede usar =countif(rango,"<=1"). (Asegúrese de incluir las comillas). Puede contar todas las observaciones que están entre 1 y -1 (inclusive) usando La función de Excel =count() cuenta todas las entradas en un rango. Por lo tanto, puede determinar la proporción de observaciones menores o iguales a 1 usando countif(rango,"<=1")/count(rango).*

*=countif(rango,"<=1") - countif(rango,"<-1").*

e) Refine su análisis completando la siguiente tabla para los datos de libras vendidas

A table of data with text

Description automatically generated with medium confidence

f) ¿Qué tan bien se ajustan los datos de libras de material vendido a la distribución normal (en forma de campana)? Justifique su respuesta con los resultados de los apartados a) a e). (Entiendo que no tiene un estándar para evaluar el ajuste; simplemente use su criterio).

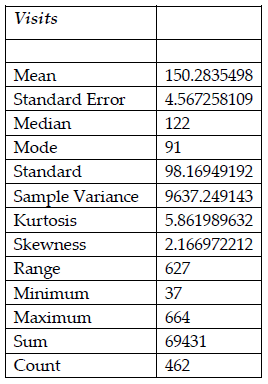
g) Determine la asimetría y la curtosis para los datos de libras vendidas. ¿Son estos valores consistentes con su análisis de los datos de libras de material vendido?

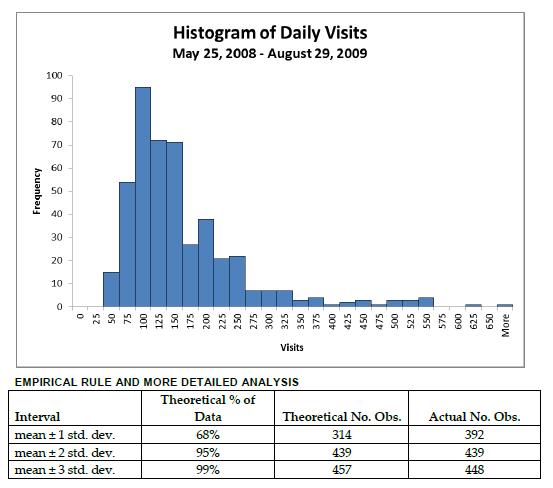
*Consejo de Excel: Asimetría y curtosis. En Excel, estos valores se pueden obtener mediante las funciones =skew() y =kurt(). También se incluyen como parte de la salida cuando se determinan las estadísticas descriptivas para un rango de datos.*

9) Como parte del análisis, también se recopiló el número de visitas diarias al sitio web de control de calidad durante el período del 25 de mayo de 2008 al 29 de agosto de 2009. El material a continuación es el resultado de los datos de visitas diarias utilizando el mismo análisis que realizó en el problema anterior para los datos de libras de material vendidas por semana. Su tarea en este problema es usar los dos conjuntos de resultados para escribir uno o dos párrafos comparando la distribución de los datos de libras vendidas con la de los datos de visitas diarias. Es decir, ¿es uno más "normal" que el otro? ¿Cómo lo sabe? (Tenga en cuenta que los datos de visitas diarias se proporcionan en la hoja de cálculo de datos de Quality Alloys; sin embargo, puede confiar aquí únicamente en el resultado a continuación).

RESULTADO DEL ANÁLISIS PARA EL NÚMERO DE VISITAS DIARIAS AL SITIO WEB DE CONTROL DE CALIDAD DURANTE EL PERÍODO DEL 25 DE MAYO DE 2008 AL 29 DE AGOSTO DE 2009

Estadísticas de resumen utilizando el comando Datos/Análisis de datos/Estadísticas descriptivas de Excel:





A table of numbers and letters

Description automatically generated

Los valores de sesgo y curtosis se proporcionan en los cálculos de estadísticas descriptivas, arriba.

RESUMEN DE DATOS/RESUMEN GRÁFICO DE DATOS

Los datos de la hoja de trabajo de Demografía se incluyen para (a) proporcionarle una idea más completa del tipo de datos web que se pueden recopilar y se recopilan, y (b) brindarle una imagen más completa de los clientes de QA y su sitio web.

10) Represente cada conjunto de datos gráficamente. En cada caso, escriba una o dos oraciones que reflejen la(s) conclusión(es) principal(es) a la que llegue