

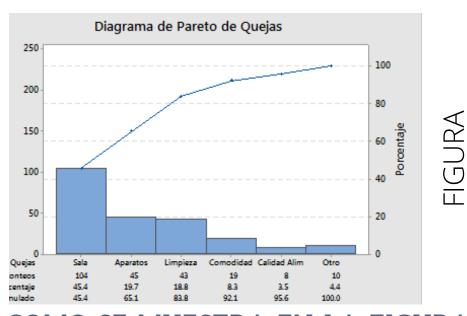
ANÁLISIS DE RESEÑAS DE HOTELES

DELGADO LOZANO O.M TREVIÑO HERNANDEZ J.P



INTRODUCCIÓN

A QUIEN NO LE GUSTARÍA PODER VIAJAR Y LLEVAR LA COMODIDAD Y DESCANSO QUE NOS OFRECE NUESTRO HOGAR, PERO COMO NO PODEMOS LLEVARNOS NUESTRO HOGAR NECESITAMOS QUEDARNOS EN UN HOTEL PERO ESO NO SIGNIFICA QUE EL HOTEL NO SE PUEDA SENTIR COMO NUESTRA CASA.



COMO SE MUESTRA EN LA FIGURA 1 EXISTEN DIFERENTES QUEJAS AL HOSPEDARTE EN UN HOTEL.

ES POR ESO QUE DECIDIMOS HACER EL ANÁLISIS DE ESTA BASE **DATOS PARA PODERLE** BRINDAR AL CONSUMIDOR Y AL OTRA VISIÓN HOTELERO Α DEL RATING **APARTE** DE COMO **ESTRELLAS** ES EL SENTIMIENTO QUE GENERA EN LOS HUÉSPEDES AL HOSPEDARSE EN ALGÚN HOTEL, Y TAMBIÉN EL HUÉSPED TENGA UNA VISIÓN DE QUE PUEDE **ESPERAR** LO EMOCIONALMENTE AL ELEGIR UN HOTEL

RECURSOS Y SOFTWARE







OBJETIVO

PREDECIR EL SENTIMIENTO PRODUCIDO POR ESTADÍA EN UN HOTEL DEL ANÁLISIS DE SUS RESEÑAS

SECUNDARIO

- CONOCER LAS PALABRAS MAS USADAS
- CONOCER EL NUMERO DE HOTELES CON 5,4,3,2,1, ESTRELLAS

METODOLOGÍA

LA BASE DE DATOS CUENTA CON DOS COLUMNAS LA PRIMERA HACE REFERENCIA A LAS RESEÑAS DE LOS USUARIOS O HUÉSPEDES Y DONDE DEBEN HACER UNA DESCRIPCIÓN DE COMO ESTUVO SU CUARTO, EL SERVICIO, LAS INSTALACIONES...ETC.

LA SEGUNDA COLUMNA SE REFIERE A UNA CALIFICACIÓN EN BASE A ESTRELLAS DONDE 5 ESTRELLAS ES UNA REFERENCIA DE EXCELENCIA Y 1 ESTRELLAS DE MUY MALO.

TAMBIÉN LA BASE DE DATOS TIENE 20491 FILAS Y COMO SE MENCIONO 2 COLUMNAS, LO BUENO DE ESTA BASE DE DATOS ES QUE NO CUENTA CON DATOS NULOS LO QUE NOS FACILITA MUCHO A LA HORA DE LA MANIPULACIÓN DE DATOS ADEMAS DE TENER MAS INFORMACIÓN PARA SU ANÁLISIS.

COMO NUESTRO OBJETIVO ES EL ANÁLISIS DE LAS RESEÑAS LO QUE HICIMOS FUE QUITAR AQUELLAS COSAS QUE NO FUERA PALABRAS COMO NÚMEROS O SIGNOS ESPECIALES, DESPUÉS LAS PALABRAS QUE ESTÉN ENTRE a-z y A-Z LAS ACOMODAMOS EN UN VECTOR SEPARADAS UNA PALABRA POR POSICIONO LUEGO POR MEDIO DEL PAQUETE STOPWORD NLTK (Natural Language DE LA LIBRERÍA Toolkit), QUE NOS AYUDA AL ANÁLISIS DE LAS PALABRAS SE DIVIDEN EN 3 CATEGORÍAS QUE SON MALO NEUTRAL BUENO Y DEPENDIENDO DE EN QUE CATEGORÍA ENTRE LAS PALABRAS SE LES ASIGNA UN VALOR A LAS PALABRAS, **TÉCNICA** GRACIAS A LA **ESTO** CLASIFICACIÓN, Y EN BASE SE SUMA EL PESO DE LAS PALABRAS PARA PODER DECIR SI UNA RESEÑA ES POSITIVA AL TENER MAS PALABRAS POSITIVAS O NEGATIVA. COMO SE VE EN LA FIGURA 2 DONDE LA COLUMNA 0 EN EJE DE LAS X SON LAS RESEÑAS NEGATIVAS LA COLUMNA 1 SON LAS RESEÑAS NEUTRALES Y LAS DE LA COLUMNA 1 SON LAS RESEÑAS **POSITIVAS**

RESULTADOS





IGURA 4



FIGURA 5

FIGURA 3: NUMERO DE RESEÑAS CON RATING 1,2,3,4,5 ESTRELLAS.

FIGURA 4: PALABRAS MAS USADAS EN LAS RESEÑAS CON RATING DE 5 ESTRELLAS.

FIGURA 5: PALABRAS MAS USADAS EN LAS RESEÑAS CON RATING DE 1 ESTRELLAS,

CONCLUSIÓN

DESPUÉS DE REALIZAR EL ANÁLISIS DE LOS DATOS QUE LAS RESEÑAS CON 5 ESTRELLAS COMO LAS DE 4 ESTRELLAS USARON PALABRAS MUY PARECIDAS COMO CUARTO, SERVICIO DÁNDONOS A ENTENDER QUE SON LAS COSAS EN LAS QUE MAS SE FIJA LA GENTE TANTO PARA LO BUENO COMO PARA LO MALO,

METAS FUTURAS: PODER HACER EL DESARROLLO DE UN BOT QUE SE LE PROPORCIONE A LOS HOTELES Y ESTE REVISANDO LAS RESEÑAS PARA EL ANÁLISIS Y PODER DETECTAR MAS RÁPIDO CUANDO UNA RESEÑA MALA APARECE Y PODER HACER EL ESTUDIO DE LA MISMA BUSCANDO ÁREAS DE OPORTUNIDAD PARA MEJOR A LOS HOTELES.

QUIERES SABER
MAS ESCANEA EL
QR

