**Select the representative district**

use hotel

go

declare @text nvarchar(1000)

declare @start int

declare @n int -- ngram

declare @len int

declare @count int

declare @ch nvarchar(10)

set @n=2

if OBJECT\_ID('tempdb.dbo.#t')is not null

drop table #t

create table #t(ch nvarchar(10)collate chinese\_prc\_ci\_as null, frq int)

CREATE INDEX idx\_t ON #t(ch)

declare curr cursor local for

select 地址 from houst\_data$

open curr

fetch next from curr into @text

while @@fetch\_status=0

begin

set @len=len(@text)

set @start=1

while @start<@len

begin

set @ch = substring(@text,@start,@n)

insert into #t(ch,frq) values(@ch,1)

set @start=@start+1

end

fetch next from curr into @text

end

close curr

go

select ch, count(\*)as frq from #t

group by ch

order by frq desc

**Create result table**

create table hotel\_data(酒店类型 char (20),平均最低价格 int,最低价格 int ,最高价格 int , 酒店数量 int)

**Insert data**

insert into hotel\_data

select '朝阳区' 行政区,'商务酒店' 酒店类型,AVG(最低价格) 平均最低价格,MIN(最低价格) 最低价格,MAX(最低价格) 最高价格,COUNT(distinct(酒店名)) 酒店数量,VAR(最低价格) 价格方差 from dbo.houst\_data$ where 地址 like '%朝阳区%' and rtrim(酒店名) like '%商务酒店%'

insert into hotel\_data

select '朝阳区' 行政区,'公寓' 酒店类型,AVG(最低价格) 平均最低价格,MIN(最低价格) 最低价格,MAX(最低价格) 最高价格,COUNT(distinct(酒店名)) 酒店数量,VAR(最低价格) 价格方差 from dbo.houst\_data$

where rtrim(酒店名) like '%公寓%' and 地址 like '%朝阳区%'

insert into hotel\_data

select '朝阳区' 行政区,'高级酒店' 酒店类型,AVG(最低价格) 平均最低价格,MIN(最低价格) 最低价格,MAX(最低价格) 最高价格,COUNT(distinct(酒店名)) 酒店数量,VAR(最低价格) 价格方差 from dbo.houst\_data$

where rtrim(酒店名) like '%酒店%' and rtrim(酒店名) not like '%[便捷快连锁速汉庭联盟布丁公寓商务]%' and 地址 like '%朝阳区%' and 最低价格>800

insert into hotel\_data

select '朝阳区' 行政区,'普通酒店' 酒店类型,AVG(最低价格) 平均最低价格,MIN(最低价格) 最低价格,MAX(最低价格) 最高价格,COUNT(distinct(酒店名)) 酒店数量,VAR(最低价格) 价格方差 from dbo.houst\_data$

where rtrim(酒店名) like '%酒店%' and rtrim(酒店名) not like '%[便捷快连锁速汉庭联盟布丁公寓商务]%' and 地址 like '%朝阳区%' and 最低价格<=800

insert into hotel\_data

select '朝阳区' 行政区,'招待所' 酒店类型,AVG(最低价格) 平均最低价格,MIN(最低价格) 最低价格,MAX(最低价格) 最高价格,COUNT(distinct(酒店名)) 酒店数量,VAR(最低价格) 价格方差

from dbo.houst\_data$

where rtrim(酒店名) like '%招待所%' and 地址 like '%朝阳区%'

insert into hotel\_data

select '朝阳区' 行政区,'宾馆' 酒店类型,AVG(最低价格) 平均最低价格,MIN(最低价格) 最低价格,MAX(最低价格) 最高价格,COUNT(distinct(酒店名)) 酒店数量,VAR(最低价格) 价格方差 from dbo.houst\_data$

where rtrim(酒店名) like '%宾馆%' and 地址 like '%朝阳区%'

insert into hotel\_data

select '朝阳区' 行政区,'大厦' 酒店类型,AVG(最低价格) 平均最低价格,MIN(最低价格) 最低价格,MAX(最低价格) 最高价格,COUNT(distinct(酒店名)) 酒店数量,VAR(最低价格) 价格方差 from dbo.houst\_data$

where rtrim(酒店名) like '%大厦%' and 地址 like '%朝阳区%'

insert into hotel\_data

select '朝阳区' 行政区,'四合院' 酒店类型,AVG(最低价格) 平均最低价格,MIN(最低价格) 最低价格,MAX(最低价格) 最高价格,COUNT(distinct(酒店名)) 酒店数量,VAR(最低价格) 价格方差 from dbo.houst\_data$

where rtrim(酒店名) like '%四合院%' and 地址 like '%朝阳区%'

insert into hotel\_data

select '朝阳区' 行政区,'大院' 酒店类型,AVG(最低价格) 平均最低价格,MIN(最低价格) 最低价格,MAX(最低价格) 最高价格,COUNT(distinct(酒店名)) 酒店数量,VAR(最低价格) 价格方差 from dbo.houst\_data$

where rtrim(酒店名) like '%大院%' and 地址 like '%朝阳区%' and rtrim(酒店名) like '%民俗院%' and 地址 like '%朝阳区%'

insert into hotel\_data

select '朝阳区' 行政区,'农家院' 酒店类型,AVG(最低价格) 平均最低价格,MIN(最低价格) 最低价格,MAX(最低价格) 最高价格,COUNT(distinct(酒店名)) 酒店数量,VAR(最低价格) 价格方差 from dbo.houst\_data$

where rtrim(酒店名) like '%农家%' and 地址 like '%朝阳区%'

insert into hotel\_data

select '朝阳区' 行政区,'客栈' 酒店类型,AVG(最低价格) 平均最低价格,MIN(最低价格) 最低价格,MAX(最低价格) 最高价格,COUNT(distinct(酒店名)) 酒店数量,VAR(最低价格) 价格方差 from dbo.houst\_data$

where rtrim(酒店名) like '%客栈%' and 地址 like '%朝阳区%' or rtrim(酒店名) like '%青年旅社%' and 地址 like '%朝阳区%'

or rtrim(酒店名) like '%驿站%' and 地址 like '%朝阳区%'

insert into hotel\_data

select '朝阳区' 行政区,'度假村' 酒店类型,AVG(最低价格) 平均最低价格,MIN(最低价格) 最低价格,MAX(最低价格) 最高价格,COUNT(distinct(酒店名)) 酒店数量,VAR(最低价格) 价格方差 from dbo.houst\_data$

where rtrim(酒店名) like '%度假%' and 地址 like '%朝阳区%' or rtrim(酒店名) like '%山庄%' and 地址 like '%朝阳区%'

or rtrim(酒店名) like '%庄园%' and 地址 like '%朝阳区%'

insert into hotel\_data

select '朝阳区' 行政区,'会所' 酒店类型,AVG(最低价格) 平均最低价格,MIN(最低价格) 最低价格,MAX(最低价格) 最高价格,COUNT(distinct(酒店名)) 酒店数量,VAR(最低价格) 价格方差 from dbo.houst\_data$

where rtrim(酒店名) like '%会所%' and 地址 like '%朝阳区%' or rtrim(酒店名) like '%会馆%' and 地址 like '%朝阳区%'

or rtrim(酒店名) like '%公馆%' and 地址 like '%朝阳区%'

or rtrim(酒店名) like '%俱乐部%' and 地址 like '%朝阳区%'

insert into hotel\_data

select '朝阳区' 行政区,'饭店' 酒店类型,AVG(最低价格) 平均最低价格,MIN(最低价格) 最低价格,MAX(最低价格) 最高价格,COUNT(distinct(酒店名)) 酒店数量,VAR(最低价格) 价格方差 from dbo.houst\_data$

where rtrim(酒店名) like '%饭%' and 地址 like '%朝阳区%'

insert into hotel\_data

select '朝阳区' 行政区,'旅馆' 酒店类型,AVG(最低价格) 平均最低价格,MIN(最低价格) 最低价格,MAX(最低价格) 最高价格,COUNT(distinct(酒店名)) 酒店数量,VAR(最低价格) 价格方差 from dbo.houst\_data$

where rtrim(酒店名) like '%旅%' and rtrim(酒店名) not like '%青年旅%' and 地址 like '%朝阳区%'

insert into hotel\_data

select '朝阳区' 行政区,'快捷酒店' 酒店类型,AVG(最低价格) 平均最低价格,MIN(最低价格) 最低价格,MAX(最低价格) 最高价格,COUNT(distinct(酒店名)) 酒店数量,VAR(最低价格) 价格方差 from dbo.houst\_data$

where rtrim(酒店名) like '%速%'and 地址 like '%朝阳区%'

or rtrim(酒店名) like '%汉庭%' and 地址 like '%朝阳区%'

or rtrim(酒店名) like '%布丁%' and 地址 like '%朝阳区%'

or rtrim(酒店名) like '%锦江之星%' and 地址 like '%朝阳区%'

or rtrim(酒店名) like '%莫泰%' and 地址 like '%朝阳区%'

or rtrim(酒店名) like '%格林豪泰%' and 地址 like '%朝阳区%'

or rtrim(酒店名) like '%便捷%' and 地址 like '%朝阳区%'

or rtrim(酒店名) like '%快捷%' and 地址 like '%朝阳区%'

or rtrim(酒店名) like '%联盟%' and 地址 like '%朝阳区%'

or rtrim(酒店名) like '%连锁%' and 地址 like '%朝阳区%'

**Explore the relationship between the price and the name**

select top 10 最低价格,酒店名,地址 from dbo.houst\_data$ order by 最低价格 desc

select top 10 最低价格,酒店名,地址 from dbo.houst\_data$ order by 最低价格

**Frequency of two-word in the price<200 group**

use hotel

go

declare @text nvarchar(1000)

declare @start int

declare @n int -- ngram

declare @len int

declare @count int

declare @ch nvarchar(10)

set @n=2

if OBJECT\_ID('tempdb.dbo.#t')is not null

drop table #t

create table #t(ch nvarchar(10)collate chinese\_prc\_ci\_as null, frq int)

CREATE INDEX idx\_t ON #t(ch)

declare curr cursor local for

select 酒店名 from houst\_data$ where 最低价格 <200

open curr

fetch next from curr into @text

while @@fetch\_status=0

begin

set @len=len(@text)

set @start=1

while @start<@len

begin

set @ch = substring(@text,@start,@n)

insert into #t(ch,frq) values(@ch,1)

set @start=@start+1

end

fetch next from curr into @text

end

close curr

go

select ch, count(\*)as frq from #t

group by ch

order by frq desc

**Frequency of two-word in the price between 200 and 500 group**

use hotel

go

declare @text nvarchar(1000)

declare @start int

declare @n int -- ngram

declare @len int

declare @count int

declare @ch nvarchar(10)

set @n=2

if OBJECT\_ID('tempdb.dbo.#t')is not null

drop table #t

create table #t(ch nvarchar(10)collate chinese\_prc\_ci\_as null, frq int)

CREATE INDEX idx\_t ON #t(ch)

declare curr cursor local for

select 酒店名 from houst\_data$ where 最低价格 between 200 and 500

open curr

fetch next from curr into @text

while @@fetch\_status=0

begin

set @len=len(@text)

set @start=1

while @start<@len

begin

set @ch = substring(@text,@start,@n)

insert into #t(ch,frq) values(@ch,1)

set @start=@start+1

end

fetch next from curr into @text

end

close curr

go

select ch, count(\*)as frq from #t

group by ch

order by frq desc

**Frequency of two-word in the price between 500 and 1000 group**

use hotel

go

declare @text nvarchar(1000)

declare @start int

declare @n int -- ngram

declare @len int

declare @count int

declare @ch nvarchar(10)

set @n=2

if OBJECT\_ID('tempdb.dbo.#t')is not null

drop table #t

create table #t(ch nvarchar(10)collate chinese\_prc\_ci\_as null, frq int)

CREATE INDEX idx\_t ON #t(ch)

declare curr cursor local for

select 酒店名 from houst\_data$ where 最低价格 between 500 and 1000

open curr

fetch next from curr into @text

while @@fetch\_status=0

begin

set @len=len(@text)

set @start=1

while @start<@len

begin

set @ch = substring(@text,@start,@n)

insert into #t(ch,frq) values(@ch,1)

set @start=@start+1

end

fetch next from curr into @text

end

close curr

go

select ch, count(\*)as frq from #t

group by ch

order by frq desc

use hotel

go

declare @text nvarchar(1000)

declare @start int

declare @n int -- ngram

declare @len int

declare @count int

declare @ch nvarchar(10)

set @n=2

**Frequency of two-word in the price>1000 group**

if OBJECT\_ID('tempdb.dbo.#t')is not null

drop table #t

create table #t(ch nvarchar(10)collate chinese\_prc\_ci\_as null, frq int)

CREATE INDEX idx\_t ON #t(ch)

declare curr cursor local for

select 酒店名 from houst\_data$ where 最低价格 > 1000

open curr

fetch next from curr into @text

while @@fetch\_status=0

begin

set @len=len(@text)

set @start=1

while @start<@len

begin

set @ch = substring(@text,@start,@n)

insert into #t(ch,frq) values(@ch,1)

set @start=@start+1

end

fetch next from curr into @text

end

close curr

go

select ch, count(\*)as frq from #t

group by ch

order by frq desc