```sql
--Create a view with calculated state averages
CREATE VIEW vwStateAvg1
AS
SELECT  avg(F.HCAP) as 'Average',
        f.[State],
        f.[Measure Start Date],
        f.[HCAHPS Measure ID],
        f.[HCAHPS Question]
FROM
    (SELECT     CAST(RH1.[HCAHPS Answer Percent] as integer) as 'HCAP',
                h.[State],
                RH1.[Measure Start Date],
                RH1.[HCAHPS Measure ID],
                m.[HCAHPS Question]
    FROM    [dbo].[HCAHPS_RESULT_HOSPITAL] AS RH1
        JOIN [dbo].[HCAHPS_HOSPITAL] AS H
            ON h.[Provider ID] = RH1.[Provider ID]
        JOIN [dbo].[HCAHPS_MEASURE] AS M
            ON RH1.[HCAHPS Measure ID] = M.[HCAHPS Measure ID]
    WHERE isnumeric(RH1.[HCAHPS Answer Percent])=1) as F
GROUP BY  f.[State], f.[Measure Start Date], f.[HCAHPS Measure ID], [HCAHPS Question];
GO

--Create table with previous measure date, calculated and given averages, and
  differences

CREATE VIEW vwPrevMeasStartDate_StateAvg
AS
SELECT  SA.[STATE],
        SA.[HCAHPS Measure ID],
        SA.[HCAHPS Question],
        SA.[Measure Start Date],
        SA.[AVERAGE],                       --Calculated average
        (   SELECT TOP 1 [Measure Start Date]
            FROM [dbo].[HCAHPS_MEASURE_RANGE] AS MR
            WHERE MR.[Measure Start Date] < SA.[Measure Start Date]
            ORDER BY MR.[Measure Start Date] DESC
        ) AS PrevMeasStartDt,
        (   SELECT [HCAHPS Answer Percent]
            FROM HCAHPS_RESULT_STATE HRS
            WHERE HRS.[HCAHPS Measure ID]=sa.[HCAHPS Measure ID]
                and hrs.State=SA.State
                and HRS.[Measure Start Date]=sa.[Measure Start Date]
        )as GivenAvg,                       --Average given in State Release
        (   SELECT [HCAHPS Answer Percent]
            FROM HCAHPS_RESULT_STATE HRS
            WHERE HRS.[HCAHPS Measure ID]=sa.[HCAHPS Measure ID]
                and hrs.State=SA.State
                and HRS.[Measure Start Date]=sa.[Measure Start Date]
                and ISNUMERIC([HCAHPS Answer Percent])=1
        )-Average as diff                   --Difference between the given and
          calculated avg
```

```sql
FROM      [dbo].[vwStateAvg1] AS SA
        JOIN [dbo].[HCAHPS_MEASURE] AS M
            ON SA.[HCAHPS Measure ID] = M.[HCAHPS Measure ID];
GO

-- Create final view with above views
CREATE VIEW vwStateAvgPerHospResults
AS
SELECT [STATE],
    [HCAHPS Measure ID],
    [HCAHPS Question],
    [Measure Start Date],
    Average,
    PrevMeasStartDt,
(    SELECT [Average]
            FROM  [vwStateAvg1] as SA2
            WHERE   SA2.[state] =SA1.[STATE] AND
                    SA2.[HCAHPS Measure ID] = SA1.[HCAHPS Measure ID] AND
                    SA2.[Measure Start Date] = SA1.PrevMeasStartDt
        ) AS PrevSTAvg
FROM [vwPrevMeasStartDate_StateAvg] as SA1;
GO

CREATE VIEW vwStateComparisions
AS
SELECT  [STATE]
        ,[HCAHPS Measure ID]
        ,[HCAHPS Question]
        ,[Measure Start Date]
        , Average as CalcAvg
        , GivenAvg
        , diff
        ,   (CASE WHEN diff < -1 or diff >1 THEN 1
              WHEN diff IS NULL THEN NULL
              ELSE 0
                END
            ) AS DiffExceeds1    --1 when difference between Given and Calc Avg exceeds⏎
               1, else 0
        ,PrevMeasStartDt
        ,   (   SELECT [Average]
                FROM  [vwStateAvg1] as SA2
                WHERE   SA2.[state] =SA1.[STATE] AND
                        SA2.[HCAHPS Measure ID] = SA1.[HCAHPS Measure ID] AND
                        SA2.[Measure Start Date] = SA1.PrevMeasStartDt
            ) AS PrevSTAvg       --Calculated average of the previous start date
        ,   (CASE WHEN Average >
                    (SELECT [Average]
                    FROM  [vwStateAvg1] as SA2
                    WHERE   SA2.[state] =SA1.[STATE] AND
                            SA2.[HCAHPS Measure ID] = SA1.[HCAHPS Measure ID] AND
                            SA2.[Measure Start Date] = SA1.PrevMeasStartDt)
                THEN 1
```

```sql
                    WHEN Average IS NULL THEN NULL
                        ELSE 0
                    END
                ) AS ScoreIncreased       --indicates whether (1) or not (0) the average  ⇀
                    increased from the previous measure
FROM [vwPrevMeasStartDate_StateAvg] as SA1
;
GO


--Create view calculating national averages
CREATE VIEW vwNatAvg
AS
SELECT AVG(F.HCAP) as 'Average', f.[Measure Start Date], f.[HCAHPS Measure ID]
FROM
    (SELECT CAST(RH1.[HCAHPS Answer Percent] as integer) as 'HCAP', RH1.[Measure Start⇀
        Date],RH1.[HCAHPS Measure ID]
    FROM    [dbo].[HCAHPS_RESULT_HOSPITAL] AS RH1
            JOIN [dbo].[HCAHPS_MEASURE] AS M
                ON RH1.[HCAHPS Measure ID] = M.[HCAHPS Measure ID]
    where isnumeric(RH1.[HCAHPS Answer Percent])=1
    ) as F
GROUP BY   f.[Measure Start Date], f.[HCAHPS Measure ID]
;
GO


-- View with previous measure dates, given averages, and difference btwn given and   ⇀
  calc averages
CREATE VIEW vwPrevMeasStartDate_NatAvg
AS
SELECT  NA.[HCAHPS Measure ID],
        NA.[Measure Start Date],
        M.[HCAHPS Question],
        NA.[AVERAGE],
        (   SELECT TOP 1 [Measure Start Date]
            FROM [dbo].[HCAHPS_MEASURE_RANGE] AS MR
            WHERE MR.[Measure Start Date] < NA.[Measure Start Date]
            ORDER BY MR.[Measure Start Date] DESC
        ) AS PrevMeasStartDt,
        (   SELECT [HCAHPS Answer Percent]
            FROM HCAHPS_RESULT_NATION HRS
            WHERE HRS.[HCAHPS Measure ID]=Na.[HCAHPS Measure ID]
                and HRS.[Measure Start Date]=na.[Measure Start Date]
                and ISNUMERIC([HCAHPS Answer Percent])=1
        )AS GivenAvg,
        (   SELECT [HCAHPS Answer Percent]
            FROM HCAHPS_RESULT_NATION HRS
            WHERE HRS.[HCAHPS Measure ID]=na.[HCAHPS Measure ID]
                and HRS.[Measure Start Date]=na.[Measure Start Date]
                and ISNUMERIC([HCAHPS Answer Percent])=1
        )-Average as diff
FROM    [dbo].[vwNatAvg] AS NA
        JOIN [dbo].[HCAHPS_MEASURE] AS M
```

```sql
                ON NA.[HCAHPS Measure ID] = M.[HCAHPS Measure ID]
                ;
GO


--Create View vwNationalAvgPerHospResults
CREATE VIEW vwNationalAvgPerHospResults
AS
SELECT   pm.[HCAHPS Measure ID]
        ,pm.[HCAHPS Question]
        ,pm.[Measure Start Date]
        ,pm.Average
        ,pm.[PrevMeasStartDt]
        ,( SELECT na.Average
             FROM vwNatAvg AS NA
             WHERE  NA.[HCAHPS Measure ID] = PM.[HCAHPS Measure ID] AND
                    NA.[Measure Start Date] = PM.PrevMeasStartDt
        ) AS PrevResult
FROM    vwPrevMeasStartDate_NatAvg AS pm
;
GO


CREATE VIEW vwNationalComparisions
AS
SELECT   NA1.[HCAHPS Measure ID],
         NA1.[HCAHPS Question],
         NA1.[Measure Start Date],
         NA1.[Average] as CalcAvg,
         NA1.GivenAvg,
         NA1.diff,
         (CASE WHEN NA1.diff < -1 or NA1.diff >1 THEN 1
              WHEN NA1.diff IS NULL THEN NULL
              ELSE 0
         END
         ) AS DiffExceeds1,
         NA1.PrevMeasStartDt,
         (   SELECT NA2.[Average]
             FROM  [vwNatAvg] as NA2
             WHERE   NA2.[HCAHPS Measure ID] = NA1.[HCAHPS Measure ID] AND
                     NA2.[Measure Start Date] = NA1.PrevMeasStartDt
         ) AS PrevNATAvg,
         (CASE WHEN NA1.Average >
                    (SELECT NA2.[Average]
                     FROM  [vwNatAvg] as NA2
                     WHERE   NA2.[HCAHPS Measure ID] = NA1.[HCAHPS Measure ID] AND
                             NA2.[Measure Start Date] = NA1.PrevMeasStartDt
                    )
              THEN 1
              WHEN NA1.Average IS NULL THEN NULL
               ELSE 0
         END
         ) AS ScoreIncreased
FROM [vwPrevMeasStartDate_NatAvg] as NA1
```

```sql
;
GO


--Shows the states with the most improved measures, over all of the data:
SELECT [STATE], Count(ScoreIncreased) as ImprovedMeasures
FROM vwStateComparisions SC
WHERE ScoreIncreased= 1
GROUP BY [STATE]
ORDER  BY ImprovedMeasures DESC
;
GO
---AK, VI tied for first with 165, then SD with 131, then ID with 127
---Worst are PR, CA, and TX with 27, 29, and 33 respectively

/*These comparisons shouldn't be taken at face value because measures
can hold both positive and negative connotations, but they show which
states have made the most changes*/


--Create view to pull the top state average for each measure
CREATE VIEW MaxScoresStateAverages
AS
    Select MAX([Average]) as 'MaxScore'
    , [HCAHPS Measure ID]
    , [HCAHPS Question]
    , [Measure Start Date]
    FROM vwStateAvgPerHospResults
    Group by  [HCAHPS Measure ID], [Measure Start Date], [HCAHPS Question]
;
GO


--Number of times states have had the highest average for a measure
SELECT Count(*) as NumofMaxMeas,sa.[State]
FROM MaxScoresStateAverages M
JOIN vwStateAvgPerHospResults SA
    ON SA.[Average]=M.MaxScore
    AND sa.[HCAHPS Measure ID]= m.[HCAHPS Measure ID]
    AND SA.[Measure Start Date] = m.[Measure Start Date]
GROUP BY sa.[state]
ORDER BY NumofMaxMeas DESC
;
GO

/*Looking at all of the data, DC has had the most instances (203) where their overall
average for a measure was the highest for a measure, compared to other state averages.
VI is second with 114*/


--Same as above, broken down by Measure Start Date, most recent shown first.
SELECT Count(*) as NumofMaxMeas,sa.[State], sa.[Measure Start Date]
FROM MaxScoresStateAverages M
```

```sql
JOIN vwStateAvgPerHospResults SA
    ON SA.[Average]=M.MaxScore
    AND sa.[HCAHPS Measure ID]= m.[HCAHPS Measure ID]
    AND SA.[Measure Start Date] = m.[Measure Start Date]
GROUP BY sa.[state],  sa.[Measure Start Date]
ORDER BY sa.[Measure Start Date] DESC, NumofMaxMeas DESC
;
GO
/*For example, DC had the most measures at the highest percentage for
that measure in the 2017-07-01 release*/


--Query to grab these measures where DC (or another state) is the max average score
SELECT sa.[state], sa.[Average], sa.[HCAHPS Measure ID], sa.[HCAHPS Question], sa.    ⮐
  [Measure Start Date]
FROM MaxScoresStateAverages M
JOIN vwStateAvgPerHospResults SA
    ON SA.[Average]=M.MaxScore
    AND sa.[HCAHPS Measure ID]= m.[HCAHPS Measure ID]
    AND SA.[Measure Start Date] = m.[Measure Start Date]
Where sa.[STATE]= 'DC'
ORDER BY [Measure Start Date] DESC;


 --Same process as above, but with the lowest scores
CREATE VIEW MinScoresStateAverages
AS
    Select MIN([Average]) as 'MinScore', [HCAHPS Measure ID], [HCAHPS Question],    ⮐
      [Measure Start Date]
    FROM vwStateAvgPerHospResults
    Group by  [HCAHPS Measure ID], [Measure Start Date], [HCAHPS Question]
;
GO

--Number of times states have had the highest average for a measure
SELECT Count(*) as NumofMinMeas, sa.[State]
FROM MinScoresStateAverages M
JOIN vwStateAvgPerHospResults SA
    ON SA.[Average]=M.MinScore
    AND sa.[HCAHPS Measure ID]= m.[HCAHPS Measure ID]
    AND SA.[Measure Start Date] = m.[Measure Start Date]
GROUP BY sa.[state]
ORDER BY NumofMinMeas DESC
;
GO
/*looking at all of the data, NE has had the most instances (180) where their overall
average for a measure was the lowest for a measure, compared to other state averages.
LA is second with 134*/


--Same as above, broken down by Measure Start Date, most recent shown first.
SELECT Count(*) as NumofMinMeas,sa.[State], sa.[Measure Start Date]
```

```sql
FROM MinScoresStateAverages M
JOIN vwStateAvgPerHospResults SA
    ON SA.[Average]=M.MinScore
    AND sa.[HCAHPS Measure ID]= m.[HCAHPS Measure ID]
    AND SA.[Measure Start Date] = m.[Measure Start Date]
GROUP BY sa.[state],  sa.[Measure Start Date]
ORDER BY sa.[Measure Start Date] DESC, NumofMinMeas DESC
;
GO
--For example, NE had the most measures with the lowest average scores (10), compared ⮐
  to other states, for the 2017/07/01 date

--Because measures hold both positive and negative connotations, these numbers help   ⮐
  only by identify extreme outliers

--Measures' average national scores
SELECT [HCAHPS Measure ID],[HCAHPS Question], AVG(CalcAvg) as 'MeasAvg'
FROM [dbo].[vwNationalComparisions]
GROUP BY [HCAHPS Measure ID], [HCAHPS Question]
ORDER BY MeasAvg Desc

--Measures that score highest, on average, on a national level, are                    ⮐
  H_COMP_6_Y_P,H_COMP_2_A_P, and H_COMP_1_A_P

SELECT [HCAHPS Measure ID],[HCAHPS Question], AVG(CalcAvg) as 'MeasAvg'
FROM [dbo].[vwNationalComparisions]
GROUP BY [HCAHPS Measure ID], [HCAHPS Question]
ORDER BY MeasAvg ASC
--Measures that score the lowest,  on average, on a national level, are                ⮐
  H_COMP_1_SN_P,H_COMP_7_D_SD,H_RECMND_DN,H_COMP_2_SN_P
--These measures all hold negative connotations (so a low score is a good thing)



SELECT * FROM vwStateComparisions
WHERE DiffExceeds1 =1
;
--795 rows out of the 30,677 rows have differences which exceed 1

--Look to see if there are any states that have repeatedly large differences
SELECT COUNT(*) as NumExceeding1Dif, [State]
FROM vwStateComparisions
WHERE DiffExceeds1 =1
GROUP BY [State]
ORDER BY NumExceeding1Dif DESC
;
--AK (111), HI (101), ND(88) are the top three
--Out of 795 where difference exceeds 1


--Check to see how many hospital results were available
SELECT COUNT(*) as NumofHospResults, C.[STATE]
```

```sql
FROM (SELECT DISTINCT RH.[Provider ID], RH.[HCAHPS Measure ID], RH.[Measure Start     ⮎
  Date], H.[State]
        FROM HCAHPS_RESULT_HOSPITAL RH
    JOIN HCAHPS_HOSPITAL H
        ON H.[provider ID]=RH.[provider ID]) as C
GROUP BY C.[STATE]
ORDER BY NumofHospResults
;
--Rankings with 1 (fewest results) and 56 (most results):
--AK (10/56), HI(9/56), ND (18/56)

SELECT DISTINCT RH.[Provider ID], RH.[HCAHPS Measure ID], RH.[Measure Start Date], H.  ⮎
  [State]
        FROM HCAHPS_RESULT_HOSPITAL RH
    JOIN HCAHPS_HOSPITAL H
        ON H.[provider ID]=RH.[provider ID]
;


--Looking into the average number of surveys completed by state
SELECT Avg(NumSurv) as AvgNumSurveys, C.[STATE]
FROM (SELECT DISTINCT RH.[Provider ID], RH.[HCAHPS Measure ID], RH.[Measure Start      ⮎
  Date], H.[State], CONVERT(Integer,[Number of Completed Surveys]) as NumSurv
        FROM HCAHPS_RESULT_HOSPITAL RH
    JOIN HCAHPS_HOSPITAL H
        ON H.[provider ID]=RH.[provider ID]
        Where ISNUMERIC([Number of Completed Surveys])=1) as C
GROUP BY C.[STATE]
ORDER BY AvgNumSurveys
;
--AK(16/53), HI (48/53), ND(19/53)
--1=fewest, 53=most number of surveys completed, on average

--Measures repeatedly exceeding difference of 1 in calculated vs given avgs
SELECT COUNT(*) as NumExceeding1Dif, [HCAHPS Measure ID],[HCAHPS Question]
FROM vwStateComparisions
WHERE DiffExceeds1 =1
GROUP BY [HCAHPS Measure ID],[HCAHPS Question]
ORDER BY NumExceeding1Dif DESC
;
--Measures with most averages exceeding difference of seem to be "Always" questions.
--Of the top 6 with the most exceeding difference of 1, 5 are "Always" questions, 4   ⮎
  are H_COMP, and all are positive.
--example: top 3:
    --(1)Patients who reported that they "Always" received help as soon as they wanted
    --(2)Patients who reported that the area around their room was "Always" quiet at   ⮎
      night
    --(3)Patients who reported that their room and bathroom were "Always" clean

--Check footnotes on exceeding 1 differences
SELECT RS.[HCAHPS Measure ID], RS.[Measure Start Date], RS.[HCAHPS Answer Percent],    ⮎
  RS.[State], RS.Footnote
FROM vwStateComparisions SC
```

```sql
    JOIN HCAHPS_RESULT_STATE RS
    ON  SC.[State]=RS.[State]
    AND SC.[HCAHPS Measure ID]=RS.[HCAHPS Measure ID]
    AND SC.[Measure Start Date]=RS.[Measure Start Date]
WHERE SC.DiffExceeds1=1
;
--Seems to be several with footnote 20: State and national averages do not include VHA⏎
    hospital data.

--Pull and count all distinct footnotes where difference exceeds 1
SELECT DISTINCT RS.Footnote, Count(RS.Footnote)
FROM vwStateComparisions SC
    JOIN HCAHPS_RESULT_STATE RS
    ON  SC.[State]=RS.[State]
    AND SC.[HCAHPS Measure ID]=RS.[HCAHPS Measure ID]
    AND SC.[Measure Start Date]=RS.[Measure Start Date]
WHERE SC.DiffExceeds1=1
GROUP BY RS.Footnote
;
--Footnote 20 is the only one present. 399 instances out of 795 have footnote 20

--Count number of #20 Footnote instances for each state
SELECT Count(Footnote) as FN20Count, RS.[State]
FROM vwStateComparisions SC
    JOIN HCAHPS_RESULT_STATE RS
        ON  SC.[State]=RS.[State]
        AND SC.[HCAHPS Measure ID]=RS.[HCAHPS Measure ID]
        AND SC.[Measure Start Date]=RS.[Measure Start Date]
WHERE SC.DiffExceeds1=1
    AND Footnote Like '20%'
GROUP BY RS.[State]
ORDER BY FN20Count DESC
;
/*Top three results are AK, HI, and ND, which are also top three results for the      ⏎
  number of
measures with differences exceeding 1 from calculated and given averages*/


--Find providers with footnotes about VHA
SELECT DISTINCT [HCAHPS Answer Percent Footnote], H.[Hospital Name],[State]
FROM HCAHPS_RESULT_HOSPITAL RH
    JOIN HCAHPS_HOSPITAL H
    On H.[Provider ID]= RH.[Provider ID]
WHERE [HCAHPS Answer Percent Footnote] LIKE  '%VHA%'
--Note all start with 'VA'

--Pull hospital names beginning with VA
SELECT DISTINCT H.[Hospital Name],[State]
FROM HCAHPS_RESULT_HOSPITAL RH
    JOIN HCAHPS_HOSPITAL H
    On H.[Provider ID]= RH.[Provider ID]
WHERE H.[Hospital Name] LIKE  'VA%'
```

```
ORDER BY [STATE]
--None listed for AK, HI, or ND
--Naming convention or footnote labeling may not be the present on the RESULT_HOSPITAL↵
    level


                    /*CONCLUSION:
State and national averages given in the resleases do not include VHA hospital data,
but they must also be included in the hospital results. This must lead to a larger
difference in the GivenAvg and CalcAvg averages. Footnote 20 is associated with this  ↵
    occurrence*/
```