



ASSISTANT MINITAB - LIVRE BLANC

Ce livre blanc fait partie d'une série de documents qui expliquent les recherches menées par les statisticiens de Minitab pour développer les méthodes et les outils de vérification des données utilisés dans l'Assistant de Minitab Statistical Software.

Plans d'expériences (DOE)

Généralités

La fonction DOE (plan d'expériences) de l'Assistant inclut un sous-ensemble de fonctions disponibles dans Minitab et utilise une procédure d'expérimentation séquentielle visant à simplifier la création et l'analyse de plans. La première phase du procédé consiste à établir des plans de criblage pour identifier les facteurs les plus importants. Nous fournissons ensuite des plans de plus haute résolution afin de détecter une éventuelle courbure et de déterminer un modèle final permettant d'identifier les paramètres de facteurs qui optimisent la réponse.

Dans cet article, nous décrivons les étapes du procédé d'expérimentation. Nous expliquons comment nous avons sélectionné les plans proposés dans l'Assistant et abordons notamment le rôle de la puissance. Nous traitons également des méthodes de détection et d'ajustement d'une courbure dans les données. Cet article décrit également la méthode utilisée pour analyser les données et identifier le modèle le plus adéquat.

Il fournit aussi des informations supplémentaires sur les contrôles des données suivants dans le rapport de l'Assistant :

- Blocs
- Données aberrantes
- Capacité de détection

Méthode

Procédé d'expérimentation séquentielle

Les fonctions de plans d'expériences de l'Assistant guident les utilisateurs au fil d'une procédure séquentielle de conception et d'analyse d'expériences visant à identifier des facteurs influents et à déterminer les paramètres de facteurs permettant d'optimiser une réponse. L'expérimentation séquentielle consiste en un ensemble de petites expériences dont les résultats orientent la suite de l'expérimentation d'une étape à la suivante. L'avantage de cette approche séquentielle est que seul un nombre réduit d'essais expérimentaux sont exécutés à chaque étape : vous évitez ainsi tout gaspillage de ressources sur des essais improductifs.

L'Assistant fournit un sous-ensemble des fonctions de DOE disponibles dans Minitab sous une forme structurée qui simplifie la création et l'analyse de plans. Les étapes de la procédure sont les suivantes :

1. Création d'un plan de criblage pour 6 à 15 facteurs.
2. Ajustement d'un modèle de criblage qui inclue les effets principaux et analyse les résultats pour détecter les facteurs les plus importants.
3. Création d'un plan de modélisation incluant 2 à 5 des facteurs les plus importants déterminés dans l'étape 2.
4. Ajustement d'un modèle linéaire qui inclue les effets principaux et les interactions à 2 facteurs, analyse des résultats et recherche d'une courbure dans la relation entre les facteurs et la réponse.
5. Si aucune courbure n'est détectée à l'étape 4, utilisez ce modèle pour identifier les paramètres de facteurs qui optimisent la réponse.
6. Si une courbure est détectée à l'étape 4, l'Assistant vous recommande d'ajouter des points pour la courbure dans le plan.
7. Ajustement d'un modèle quadratique incluant des termes au carré pour modéliser la courbure et analyser les résultats.
8. En utilisant le modèle final, identification des paramètres de facteurs qui optimisent la réponse.

Les sections suivantes fournissent des informations détaillées sur les aspects suivants des plans d'expériences (DOE) de l'Assistant :

- Plans de criblage
- Plans de modélisation
- Ajustement du modèle

Plans de criblage

En général, au début d'une procédure d'expérimentation séquentielle, vous disposez d'un nombre élevé de facteurs potentiels et vous éliminez ensuite ceux dont l'effet sur la réponse est minime. Les plans de criblage sont des plans expérimentaux visant à identifier les facteurs les plus influents au sein d'un ensemble plus large. L'Assistant vous propose des plans de criblage pour 6 à 15 facteurs.

Type de plan

Les plans de criblage de l'Assistant sont des plans de Plackett-Burman, un type particulier de plan à 2 niveaux et de résolution III. Les plans de Plackett-Burman présentent deux avantages principaux :

- Ils permettent d'estimer les effets principaux des facteurs avec très peu d'essais expérimentaux (aussi peu que 12). La réalisation des essais expérimentaux pouvant être coûteuse, cela rend ces modèles plus économiques.
- Le confusion entre les effets principaux et les interactions à deux facteurs n'est que partielle (ou fractionnaire). Les effets qui ne peuvent pas être évalués séparément sont dits "confondus". Dans les plans de Plackett-Burman, la confusion est considérée comme partielle car la contribution de chaque effet n'est qu'une fraction de l'importance totale de l'effet d'interaction.

Nous avons constaté que pour le criblage, l'utilisation des plans de Plackett-Burman, qui évaluent uniquement les effets principaux et pas les termes d'interaction, constituait une méthode raisonnablement acceptable. Les plans de criblage permettent d'inclure un nombre élevé de facteurs. Etant donné que chaque terme du modèle requiert au moins un essai et que le nombre de termes d'interaction augmente plus vite que celui des effets principaux, il est souvent peu pratique ou peu rentable d'ajuster un modèle avec des interactions. En outre, dans la plupart des cas, seul un petit nombre de facteurs expliquent la plupart des effets sur la réponse. L'objectif d'un plan de criblage est d'identifier ces facteurs et les plans de Plackett-Burman permettent aux utilisateurs d'identifier ces effets principaux importants. De plus, comme indiqué précédemment, la confusion entre les termes n'étant que partielle dans les plans de Plackett-Burman, il est moins probable qu'un effet principal significatif soit en fait une interaction significative à 2 facteurs.

Puissance et repli

Lorsque nous avons conçu notre catalogue de plans, notre objectif était de proposer uniquement des plans ayant une puissance suffisante. Nous avons calculé la puissance de tous les plans, puis éliminé ceux dont elle était trop faible, notamment le plan à 12 essais de Plackett-Burman pour 10 ou 11 facteurs. Pour les plans à 10 ou 11 facteurs, seul le plan

à 20 essais de Plackett-Burman est disponible. Nous avons également éliminé les plans à 16, 17 et 18 facteurs en raison de leur faible puissance et du nombre d'essais important qu'ils requièrent. Pour plus d'informations sur la puissance spécifique des plans, reportez-vous à la section Capacité de détection.

Pour les plans à 6-9 facteurs, le repli est autorisé, ce qui permet d'ajouter des essais à l'expérience et d'augmenter ainsi la précision et la puissance du plan. Dans certains cas, il peut être souhaitable d'ajouter des essais à un plan pour augmenter la probabilité de détecter des effets importants. Le repli permet d'ajouter de nouveaux essais au plan, dans lesquels une partie ou la totalité des niveaux de facteurs sont inversés (les niveaux inférieur et supérieur sont inversés). Le repli permet aussi d'éliminer la confusion partielle entre les effets principaux et les interactions à deux facteurs, ce qui réduit le biais dû à la confusion lors de l'évaluation des effets principaux. La section Capacité de détection du rapport de la fonction Créer un plan de criblage fournit des informations afin d'aider les utilisateurs à déterminer si le plan est suffisamment puissant pour détecter des effets d'une certaine importance.

Plans de modélisation

Lorsque 2 à 5 facteurs importants sont identifiés, Minitab recommande de créer un plan de modélisation afin d'obtenir un modèle permettant d'identifier les paramètres de facteurs qui optimisent la réponse.

Type de plan

Les plans de modélisation à 2 ou 5 facteurs sont des plans factoriels complets ou de résolution V. Ces plans permettent notamment d'ajuster l'ensemble des termes d'effet principal et d'interaction à 2 facteurs sans aucune confusion. Une partie ou l'ensemble des termes d'ordre supérieur (c'est-à-dire les interactions à 3 facteurs) peuvent être confondus avec les termes du modèle. Cependant, les termes d'ordre supérieur sont supposés négligeables par rapport aux termes d'effet principal et aux termes d'interaction à 2 facteurs.

Lorsque nous avons conçu notre catalogue de plans, notre objectif était de proposer uniquement des plans ayant une puissance suffisante. Par conséquent, nous avons éliminé le plan à 2 facteurs et 4 essais au profit d'un plan répliqué à 4 essais et 2 facteurs.

Points centraux et modélisation de la courbure

Les plans de modélisation de l'Assistant incluent également des points centraux pour contrôler la présence d'une courbure dans les données. Il s'agit des points pour lesquels tous les facteurs continus sont définis à mi-valeur entre les paramètres minimal et maximal. S'il n'y a pas de courbure, la réponse moyenne au point central est égale à la moyenne des réponses moyennes des facteurs lorsqu'ils sont définis sur leurs paramètres minimal et maximal (les sommets de l'espace du plan). Une courbure est détectée lorsque la moyenne des réponses moyennes aux points centraux est significativement supérieure ou inférieure à celle des facteurs lorsqu'ils sont définis sur leurs paramètres minimal et maximal.

Bien que les points centraux permettent de détecter une courbure, ils ne fournissent pas suffisamment d'informations pour la modéliser. Pour modéliser la courbure, des termes carrés sont nécessaires, ce qui requiert l'ajout de points supplémentaires dans le plan. Ces

points supplémentaires changent le plan en plan composite centré à face centrée. Il s'agit d'une forme de plan de surface de réponse qui permet l'ajustement d'un modèle quadratique incluant des effets principaux linéaires, l'ensemble des interactions à 2 facteurs et les termes carrés de la totalité des facteurs continus.

Ajustement des modèles à l'aide d'une sélection descendante

Nous avons étudié plusieurs méthodes d'ajustement des modèles et constaté que la sélection descendante avec une valeur α de 0,10 semblait la plus appropriée. Lorsque vous ajustez un modèle, Minitab inclut dans un premier temps tous les termes possibles. Puis, Minitab supprime un par un les termes les moins significatifs tout en conservant la hiérarchie du modèle. Le maintien de la hiérarchie signifie que lorsqu'un terme d'interaction est significatif, les termes linéaires des deux facteurs formant l'interaction doivent également être inclus dans le modèle. Il s'agit d'une forme de sélection descendante, visant à automatiser un procédé de sélection de modèle généralement réalisé manuellement. Dans tous les plans de la fonction DOE de l'Assistant, les termes sont indépendants ou presque (dans le cas des termes carrés). Par conséquent, la multicollinéarité, c'est à dire l'existence d'une corrélation entre des facteurs, est peu probable. La multicollinéarité peut amener les procédures pas à pas à passer à côté du meilleur modèle. L'utilisation d'une valeur $\alpha = 0,10$ au lieu de la valeur habituelle $\alpha = 0,05$ accroît également la puissance des tests, ce qui augmente la probabilité que les termes importants soient conservés dans le modèle.

Vérification des données

Blocs

Les blocs sont utilisés dans des plans expérimentaux pour minimiser le biais et l'erreur dus à des influences externes sur l'expérience, telles que des variables de bruit, des variables manquantes ou des différences dans la manière dont les essais ont été réalisés dans chaque bloc. En plaçant les essais expérimentaux effectués ensemble dans des blocs, vous pouvez déterminer s'il existe des différences entre les blocs et tenir compte de ces différences dans le modèle.

Objectif


La fonction DOE de l'Assistant vous permet également d'inclure des répliques du plan lors de la création d'un plan de modélisation et d'ajouter des points axiaux à ce dernier afin d'ajuster la courbure dans le modèle. Les répliques et les points axiaux sont souvent utilisés dans des conditions ou à des moments différents des essais du plan de base. Lorsque des essais sont effectués dans des conditions ou à des moments différents, il est recommandé de tenir compte des effets possibles de ces différentes conditions.

Méthode

Pour rendre compte des différences possibles entre les conditions expérimentales des répliques ou points axiaux et celles du plan de base, Minitab place les répliques et les points axiaux dans des blocs distincts. Plus spécifiquement, dans les plans de modélisation, Minitab place les répliques du plan de base dans des blocs distincts du modèle. Dans les plans quadratiques, Minitab place les points axiaux utilisés pour détecter une courbure dans le plan à l'intérieur d'un bloc distinct.

Les résultats

Pour maintenir une certaine cohérence avec le traitement d'autres termes du modèle, les blocs sont évalués à l'aide de la méthode d'élimination descendante. Le rapport indique si le terme du bloc est statistiquement significatif, ce qui signifie qu'il existe des différences entre les blocs. S'il existe une différence entre les blocs, essayez d'en évaluer la cause afin de détecter d'éventuelles disparités dans les conditions ou les procédures expérimentales.

Etat	Condition
	<p>Les blocs se trouvent dans le modèle final</p> <p>Les blocs sont significatifs. Etant donné que les essais dans chaque bloc sont généralement réalisés à des moments différents, cette différence significative indique que les conditions ont peut-être changé dans le temps. Cette différence peut être due à des influences externes sur l'expérience, telles que des variables de bruit, des variables manquantes qui auraient dû être incluses dans l'expérience ou des différences dans la manière dont les essais ont été réalisés dans chaque bloc. Etudiez la cause de la différence entre les blocs.</p> <p>Les blocs ne se trouvent pas dans le modèle final</p> <p>Les blocs ne sont pas significatifs. Etant donné que les essais dans chaque bloc sont généralement réalisés à différents moments, ce résultat indique qu'il n'y a pas de preuve de l'existence de différences dans les conditions expérimentales dans le temps.</p>

Données aberrantes

Dans les procédures de plans d'expériences (DOE) de l'Assistant, nous définissons les données aberrantes comme des observations dont les valeurs résiduelles normalisées sont élevées, ce qui est un indicateur courant permettant d'identifier les données aberrantes dans les procédures d'ajustement de modèle (Neter et al., 1996). Les données aberrantes pouvant avoir une forte influence sur les résultats de l'analyse, il peut être nécessaire de corriger les données pour que l'analyse soit valide.

Objectif

Nous voulions déterminer à quel point les valeurs résiduelles normalisées devaient être élevées pour signaler qu'un point de données est aberrant.



Méthode

Nous avons développé des valeurs de référence permettant d'identifier les observations aberrantes suivant la procédure standard de plans d'expériences de Minitab (Stat > DOE (plan d'expériences) > Factoriel > Analyser un plan factoriel et Stat > DOE (plan d'expériences) > Surface de réponse > Analyser un plan de surface de réponse).

Les résultats

La valeur résiduelle normalisée est égale à la valeur résiduelle, e_i , divisée par une estimation de son écart type. Habituellement, une observation est considérée comme aberrante si la valeur absolue de la valeur résiduelle normalisée est supérieure à 2. Néanmoins, cette valeur est quelque peu prudente. En général, avec des ensembles de données importants, environ 5 % de la totalité des observations répondent à ce critère du simple fait du hasard (si les erreurs sont distribuées normalement). En revanche, avec de petits ensembles de données, peu d'observations sont signalées comme aberrantes par hasard, voire aucune, et il est donc recommandé de rechercher la cause des valeurs aberrantes détectées.

Lors du test des données aberrantes, l'Assistant affiche les indicateurs d'état suivants dans le rapport :

Etat	Condition
	<p>Pas de valeurs résiduelles normalisées ≥ 2</p> <p>Il n'y a aucun point de données aberrant. Les données aberrantes peuvent avoir une grande influence sur les résultats.</p>
	<p>Une seule valeur résiduelle normalisée ≥ 2</p> <p>Un point de données présente une valeur résiduelle élevée et n'est pas correctement ajusté par le modèle. Ce point apparaît en rouge sur le rapport de diagnostic et se trouve à la ligne X de la feuille de travail. Les données aberrantes pouvant fortement influencer sur les résultats, essayez d'identifier la cause des aberrations. Corrigez les erreurs de mesure ou d'entrée des données. Réalisez à nouveau les essais associés aux causes spéciales et réexécutez l'analyse.</p> <p>Plus d'une valeur résiduelle normalisée ≥ 2</p> <p>x points de données présentent des valeurs résiduelles élevées et ne sont pas correctement ajustés par le modèle. Ces points apparaissent en rouge sur le rapport de diagnostic. Pour identifier les lignes de la feuille de travail, placez votre curseur sur un point ou utilisez la fonction de balayage de Minitab. Les données aberrantes pouvant fortement influencer sur les résultats, essayez d'identifier la cause des aberrations. Corrigez les erreurs de mesure ou d'entrée des données. Réalisez à nouveau les essais associés aux causes spéciales et réexécutez l'analyse.</p>

Capacité de détection

Lors de la réalisation d'expériences dans le cadre d'un plan, il est utile de connaître l'importance d'effet pouvant être détectée par ce plan avant de collecter les données. Si le plan n'est pas suffisamment puissant pour détecter l'importance d'effet souhaitée, il est peut-être nécessaire d'ajouter des essais. Cependant, l'ajout d'essais supplémentaires à un plan pouvant être coûteux, il est important de déterminer si une puissance supplémentaire est réellement nécessaire.

Objectif

Nous voulions fournir aux utilisateurs des informations sur l'importance d'effet pouvant être détectée par leur plan avec des niveaux de puissance de 60 % et 80 %. Nous souhaitons aussi leur indiquer l'importance d'effet pouvant être détectée dans les plans qui incluent des essais supplémentaires, lorsque cela est possible. Pour les plans de criblage à 6-9 facteurs, les utilisateurs peuvent inclure 12 ou 24 essais dans leur plan. Pour les plans de modélisation, les utilisateurs peuvent inclure des répliques de leur plan de base, augmentant ainsi le nombre d'essais total du plan.

Méthode

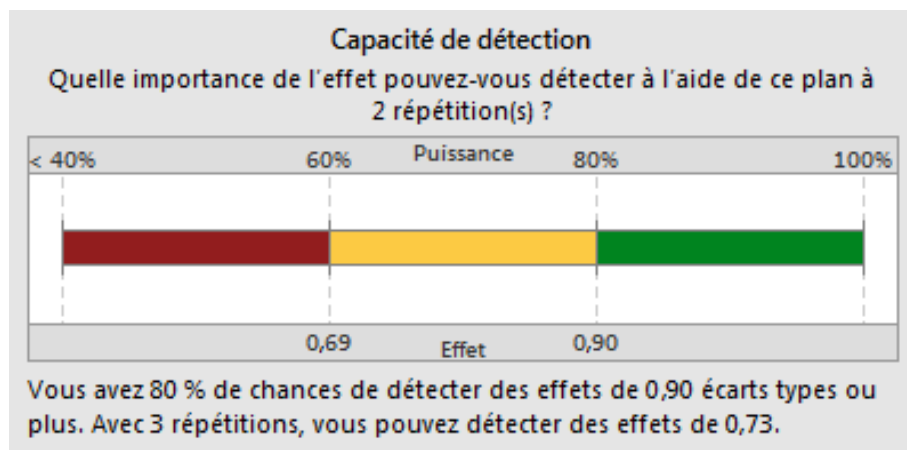
Nous avons calculé la puissance des plans de l'assistant et l'importance d'effet pouvant être détectée par chacun d'entre eux. La puissance est la probabilité de détecter que l'effet d'un facteur est statistiquement significatif. L'importance des effets est exprimée en écarts types.

Les résultats

Le rapport récapitulatif indique les importances d'effet que votre plan permet de détecter avec une puissance de 60 % et de 80 %. Dans les plans de criblage, lorsqu'un plan plus grand (replié) est disponible, le rapport indique également l'importance d'effet pouvant être détectée avec une puissance de 80 % dans le plan de plus grandes dimensions. Dans les plans de modélisation, lorsque des répliques supplémentaires sont disponibles, le rapport indique l'importance d'effet pouvant être détectée avec une puissance de 80 % grâce à l'ajout des répliques. L'utilisateur peut ainsi décider si le plan sélectionné est adapté et s'il n'est pas plus avantageux d'utiliser un plan avec plus d'essais, lorsque cela est possible.

Pour des informations plus précises sur l'importance d'effet pouvant être détectée par chaque plan avec une puissance de 60 % et de 80 %, reportez-vous à l'Annexe A.

L'image suivante est un exemple du type d'informations fournies par le rapport concernant la puissance.



Références

Neter, J., Kutner, M.H., Nachtsheim, C.J. et Wasserman, W. (1996), *Applied linear statistical models*, Chicago : Irwin.

Annexe A : capacité de détection

Nous avons calculé la puissance des plans de l'assistant et l'importance d'effet pouvant être détectée par chacun d'entre eux. La puissance est la probabilité de détecter que l'effet d'un facteur est statistiquement significatif. L'importance des effets est exprimée en écarts types.

L'importance d'effet associée à un terme du modèle est égale à deux fois le coefficient de ce terme dans l'équation réelle du modèle. Dans un modèle de criblage, l'importance d'un effet est facile à interpréter : il s'agit de la variation de la réponse moyenne lorsqu'un facteur passe de son niveau inférieur à son niveau supérieur.

Tableau 1 Le tableau suivant indique les importances d'effet pouvant être détectées par les plans de criblage disponibles dans l'Assistant.

Facteurs	Nombre d'essais	Effet avec puissance de 60 %	Effet avec une puissance de 80 %
6	12	1,27325	1,67693
6	24	0,80721	1,05805
7	12	1,3282	1,75498
7	24	0,80936	1,06092
8	12	1,43101	1,90493
8	24	0,8118	1,06420
9	12	1,68682	2,29728
9	24	0,81462	1,06797
10	20	0,919135	1,20607
11	20	0,928949	1,21945
12	20	0,941923	1,23725
13	20	0,959863	1,26206
14	20	0,986258	1,29895
15	20	1,02882	1,3594

Tableau 2 Le tableau suivant indique les importances d'effet pouvant être détectées par les plans de modélisation disponibles dans l'Assistant.

Nombre total de facteurs	Facteurs de catégorie	Répliques	Effet avec puissance de 60 %	Effet avec une puissance de 80 %
--------------------------	-----------------------	-----------	------------------------------	----------------------------------

Nombre total de facteurs	Facteurs de catégorie	Répliques	Effet avec puissance de 60 %	Effet avec une puissance de 80 %
2	0	2	1,517676	1,995488
2	0	4	0,991255	1,299356
2	0	6	0,795584	1,042572
2	1	2	1,489312	1,956272
2	1	4	0,986308	1,292747
2	1	6	0,79336	1,039623
2	2	2	1,626711	2,149402
2	2	4	1,012132	1,32742
2	2	6	0,805317	1,055524
3	0	1	1,752624	2,333048
3	0	2	1,001723	1,313394
3	0	3	0,79841	1,046325
3	0	4	0,685061	0,89768
3	0	5	0,609738	0,798946
3	1	1	1,626711	2,149402
3	1	2	0,994252	1,303368
3	1	3	0,795584	1,042572
3	1	4	0,683497	0,895612
3	1	5	0,608716	0,797597
3	2	1	1,468798	1,928128
3	2	2	0,977848	1,281481
3	2	3	0,788844	1,033647
3	2	4	0,679641	0,890522
3	2	5	0,606149	0,794214
3	3	1	3,804252	5,7928
3	3	2	1,038597	1,363392
3	3	3	0,811803	1,064195

Nombre total de facteurs	Facteurs de catégorie	Répliques	Effet avec puissance de 60 %	Effet avec une puissance de 80 %
3	3	4	0,692413	0,907434
3	3	5	0,614534	0,805288
4	0	1	1,053102	1,383293
4	0	2	0,689744	0,903887
4	0	3	0,556612	0,729334
4	0	4	0,47976	0,628615
4	0	5	0,428010	0,560802
4	1	1	1,038597	1,363392
4	1	2	0,688304	0,901977
4	1	3	0,556027	0,728562
4	1	4	0,479427	0,628176
4	1	5	0,427789	0,560511
4	2	1	1,006462	1,319772
4	2	2	0,684233	0,896585
4	2	3	0,554302	0,726288
4	2	4	0,478427	0,626861
4	2	5	0,427119	0,559631
4	3	1	0,982394	1,287529
4	3	2	0,679988	0,890980
4	3	3	0,552383	0,723762
4	3	4	0,477284	0,625358
4	3	5	0,426341	0,558609
4	4	1	1,102670	1,452267
4	4	2	0,694658	0,910421
4	4	3	0,558674	0,732059
4	4	4	0,480955	0,630190
4	4	5	0,428812	0,561858

Nombre total de facteurs	Facteurs de catégorie	Répliques	Effet avec puissance de 60 %	Effet avec une puissance de 80 %
5	0	1	1,460831	1,989497
5	0	2	0,694658	0,910421
5	0	3	0,557797	0,730899
5	0	4	0,480244	0,629252
5	0	5	0,428261	0,561133
5	1	1	1,239292	1,649714
5	1	2	0,692413	0,907434
5	1	3	0,557051	0,729913
5	1	4	0,47985	0,628733
5	1	5	0,428010	0,560802
5	2	1	1,053102	1,383293
5	2	2	0,686516	0,899606
5	2	3	0,554925	0,727108
5	2	4	0,478694	0,627212
5	2	5	0,427261	0,559817
5	3	1	0,994252	1,303368
5	3	2	0,680992	0,892303
5	3	3	0,552683	0,724156
5	3	4	0,477418	0,625533
5	3	5	0,426414	0,558704
5	4	1	0,970149	1,271267
5	4	2	0,676819	0,886805
5	4	3	0,550801	0,721681
5	4	4	0,476297	0,624062
5	4	5	0,425652	0,557704
5	5	2	0,703042	0,92162
5	5	3	0,560538	0,734525

Nombre total de facteurs	Facteurs de catégorie	Répliques	Effet avec puissance de 60 %	Effet avec une puissance de 80 %
5	5	4	0,481695	0,631166
5	5	5	0,429191	0,562356

© 2020 Minitab, LLC. All rights reserved. Minitab®, Minitab Workspace™, Companion by Minitab®, Salford Predictive Modeler®, SPM®, and the Minitab® logo are all registered trademarks of Minitab, LLC, in the United States and other countries. Additional trademarks of Minitab, LLC can be found at www.minitab.com. All other marks referenced remain the property of their respective owners.